A PREDICTIVE MODEL OF STUDENT LOAN DEFAULT AT A TWO-YEAR COMMUNITY COLLEGE

Doctoral Dissertation Research

Submitted to the Graduate Faculty of

Argosy University, Phoenix

College of Business

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Business Administration

International Business

By

Chanda Denea Brown

April 2015

A PREDICTIVE MODEL OF STUDENT LOAN DEFAULT

AT A TWO-YEAR COMMUNITY COLLEGE

A PREDICTIVE MODEL OF STUDENT LOAN DEFAULT

AT A TWO-YEAR COMMUNITY COLLEGE

Doctoral Dissertation Research

Submitted to the Faculty of

Argosy University, Phoenix

College of Business

In Partial Fulfillment

of the Requirements for the Degree of

Doctor of Business Administration

International Business

By

Chanda Denea Brown

Dissertation Committee Approval:

Chris Marcum

Chris Marcum, Ph.D. Chair

5/28/15

Date

Claudia White

Claudia White, Ph.D. Member

Edward Slover, Ph.D. Department Chair

ABSTRACT

This study explored whether a predictive model of student loan default could be developed with data from an institution's three-year cohort default rate report. The study used borrower data provided by a large two-year community college. Independent variables under investigation included total undergraduate Stafford student loan debt, total number of Stafford loan servicers, year of birth, gender, and last reported enrollment status at that institution. Two logistic regression analyses were conducted on stratified random samples to test and validate the resulting model. Descriptive statistics were calculated for the population overall, as well as generation-specific groups—millennial, generation X, and baby boomer. Results failed to develop a predictive model of student loan default. Additional research to identify other predictors of student loan repayment status would be beneficial for predicting student loan default and the development of default resolution plans.

ACKNOWLEDGEMENTS

The author would like to express sincere gratitude to committee members, Dr. Marion C. Marcum, and Dr. Claudia White for their invaluable support and guidance in the planning and implementation of this research project. The researcher would like to express appreciation to a colleague and friend, Derrick Little, for his continued support, motivation, and understanding throughout this process. The deepest appreciation is further offered to the managers and employees of the two-year community college for their participation in the study. Without their contributions of time and resources, this study would not have been possible.

DEDICATION

To my family members whose love and support have helped turn this once lifelong dream into a shared reality. To my brother Jacob, for his continued support and "motivational speeches" when all seemed lost—I thank you. Your faith in me knows no bounds, and I am lucky to call you my brother. My father Russell, I thank you for your calm demeanor to help keep things in perspective when roadblocks presented themselves. You push me to challenge myself and without your initial prompting this goal would not have been accomplished. My sister-in-law Jeanah, I thank you for your love, encouragement and support throughout this process.

To my niece Orion, please know that anything is possible if you put your mind to it. With hard work and dedication, you can accomplish your goals and exceed your expectations. To my mother Jena, who raised her children always to go after their dreams and face obstacles head on, you left the world far too soon, but I know you have been with me every step of the way and are smiling down on me.

Without all of you, this never would have been accomplished. I love you.

TABLE OF CONTENTS

Page

TABLE OF TABLES x

TABLE OF APPENDICES xi

CHAPTER ONE: INTRODUCTION 1
Purpose and Nature of Study 1
Problem Statement 3
Research Questions 3
Null Hypothesis 4
Definitions 4
Theoretical Framework 6
    Economic 7
        Human capital theory 7
        Ability to pay theory 7
        Debt burden theory 8
    Sociological 8
        Student departure theory 9
        Approach-avoidance theory 9
        General model for assessing change 10
    Determining Appropriate Theory(ies) for the Study 10
Delimitations 11
Limitations 12
Justification and Significance of Study 13

CHAPTER TWO: LITERATURE REVIEW 15
Gaps in Available Research 15
How this Research Adds to the Knowledge Base 16
Default Characteristics 17
    Borrower Demographics 17
        Ethnicity 18
        Age 18
        Gender 19
    Institutional Characteristics 19
        Type of institution 20
        Major selection and program of study 20
        Program completion 21
        College GPA 22
        Tuition and fees 22
    Socioeconomic Factors 23
        Income 23
        Debt burden 24
    State-Specific Default Borrower Research 25

Nevada 25
Florida..26
California 27
Regulatory Changes..28
Budget Control Act..28
Program Integrity..29
Amendments to the Higher Education Reauthorization Act 29
Direct Loan and Federal Family Education Loan Program (FFELP) 30
Consequences of High Cohort Default Rates 30
Pell 31
Interest Rates 31
Institutional Consequences 31
Local, State, and National Economic Effects 32
Importance of Study 33
What is a Model..34
Assessing the Merit of a Model 35
Appropriateness of Methods..36
Research Design 36
Logistic Regression Outcomes in Relation to Research Questions......................37
Data Measurements 37
Default Rate Population..38
Stratified Random Sampling 39

CHAPTER THREE: METHODOLOGY 40
Research Questions..40
Null Hypothesis 40
Research Design 40
Instrumentation..41
Population..41
Sampling..42
Standardization of Variables 43
Descriptive Analyses Procedures 43
Multivariate Analyses Procedures 44

CHAPTER FOUR: RESULTS 45
Data Review and Transformation..45
Missing Data and Recoding..46
Outliers and Linearity..46
Outliers 47
Linearity..48
Descriptive Statistics 48
Stratified Random Sampling 49
Logistic Regression Outcomes 50
Sample Group One 50
Sample Group Two..51

CHAPTER FIVE: CONCLUSIONS 54
Discussion of Results 54
Conclusion 55
Theoretical Inferences 56
    Economic 56
    Sociological 57
Implications for Practice 58
Recommendations for Future Research 59
    Millennial Generation 59
    Multi-Generational 59
    Institutional Type 60
    Institution-Specific Variable 60
    Borrower-Specific Variable 60

REFERENCES 62

APPENDICES 69

TABLE OF TABLES

Table | Page

1. Numerical Values for Variable Outcomes 43

2. Recoding of Categorical Variables 46

3. Table of Residuals Statistics to Identify Outliers 47

4. Table of Coefficients 48

5. Population and Frequency Percentage 49

6. Random Samples Stratification Values 50

7. Sample One: Variables in the Equation 51

8. Sample Two: Predictive Model Variables 52

9. Omnibus Tests of Model Coefficients 53

10. Models Summary for Variance in Loan Repayment 53

TABLE OF APPENDICES

Appendix	Page
A. 2011 Three-Year CDR Report Variables Defined	70
B. Population Demographics and Relevant Statistics	72
C. Millennial Generation Demographics and Relevant Statistics	75
D. Generation X Demographics and Relevant Statistics	78
E. Baby Boomer Demographics and Relevant Statistics	81
F. Sample One Outputs and Data Tables	84
G. Sample Two Outputs and Data Tables	89

CHAPTER ONE: INTRODUCTION

Since 2005, national student loan two-year default rates have increased—from 5.2% in 2006 to 10.0% in 2011 (Federal student aid, 2010b). While a relatively sizeable amount has been written on the characteristics of default borrowers, the vast majority of these studies are outdated (Flint, 1997; General Accounting Office, 1991; Volkwein, Cabrera, Szelest, & Napierski, 1998). The loss of viable national databases has led to an increase in institutional and state-specific student loan default research. Recent research focuses either on institutional characteristics (Barone, Steiner, & Teszler, 2005; Christman, 2000); or on state-specific defaulted borrower trends (Kypuros, 2009; Salas-Amaro, 2008). This study is intended to fill a gap in available knowledge to update the literature regarding undergraduate student loan default relevant for today's economic and social contexts.

Purpose and Nature of Study

The purpose of this study is to highlight institutional trends in borrowing repayment patterns and identify target populations who fit the category of being at-risk for student loan default. The study is not intended to develop a model that denies components of the borrower population opportunities to receive funding, but to identify those most at-risk to guide federal policy and institutional, national, and loan-servicer-based default management practices.

Appropriate for a different historical context, the applicability of previous studies focused on highlighting trends on defaulted student loan borrowers leaves much to question when compared to today's defaulted borrowers. A lack of viable datasets is largely responsible for the paradigm shift in default research from a national scope to institutional and state-centric approaches. The most vigorous studies on student loan default utilizing a national dataset occurred during the late 1980s and into the mid-1990s (Gross et al., 2009). Due to the availability of data at the time national studies were conducted, Flint (1997) and Volkwein et al. (1998) focused collectively on baby boomer and "X" generation repayment patterns, born between 1946 through 1964 and 1965 through 1981 respectively.

Findings from this study can be used for future research to compare previous findings of default trends in baby boomer and "X" generational borrowers to evaluate whether they continue to hold true in today's context. Over the last two decades, a new generation, the millennials, born between 1982 through 2002, has entered into college and graduated, leading the way to questions of what, if any, differences exist in borrower repayment trends over time and between the generations. This research adds knowledge to the field in three ways:

- to update the institutional model of defaulted borrower characteristics for undergraduate academic pursuit at a large two-year public community college;
- to provide an overview of borrower trends for the institution's 2011 three-year cohort default rate (CDR) population; and
- to provide an overview of borrower trends for the baby boomer, generation X, and millennial generations.

The study examines characteristics of defaulted Stafford loan borrowers controlling for year of birth, gender, last reported enrollment status at that institution, number of loan servicers, and total unpaid Stafford loan debt. The research is exploratory in nature and utilized the logistic regression statistical technique to investigate whether relationships exist between the identified factors. All analyses used data obtained from the host institution's National Student Loan Database System three-year CDR report. This report listed borrowers who entered repayment beginning in 2011 and followed them over a three-year span.

Problem Statement

Total outstanding student loan debt guaranteed by the United States (U.S.) Government exceeds one trillion U.S. dollars, with researchers speculating the industry cannot fail without substantial impacts on the U.S. economy (Chopra, 2012). Given the size of the industry and subsequent high costs to stakeholders, a need is present to develop a model that adequately identifies individuals who are at future risk of delinquency and potential default. Identifying those most likely to default can aid policymakers, lenders, loan servicers, and institutions in developing and targeting specific populations in their default management strategy.

Research Questions

1. Can loan repayment status (default or non-default) be correctly predicted for an institutional model from knowledge of total undergraduate Stafford student loan debt, total number of Stafford loan servicers, year of birth, gender, and last reported enrollment status at that institution?
2. If loan repayment status can be predicted correctly, which independent variables are central to the prediction of that status?
3. How many defaulted borrowers are classified correctly?

Null Hypothesis

Loan repayment status (default or non-default) cannot be predicted correctly for an institutional model from knowledge of total undergraduate Stafford student loan debt, total number of Stafford loan servicers, year of birth, gender, and last reported enrollment status at that institution.

Definitions

Baby-boomer is defined as individuals born between 1946-1965 (Catalyst, 2012; Conference Board of Canada, 2009).

Delinquency is defined as a student who is late 30 days or more on his or her Stafford loan repayments (Cunningham & Kienzel, 2011).

Default is defined as a student who has failed to make payments on his or her student loans for 270-360 days (Kantrowitz, 2009).

Default management program is defined as a program designed to deliver timely repayment communications, options for delinquent students, and solutions for defaulted borrowers.

Direct Loan (subsidized and unsubsidized) is defined as a federally guaranteed student loan with fixed interest rates under Title IV Federal Aid policy.

ED is defined as the United Stated Department of Education.

Enrollment status is defined as the borrowers last recorded enrollment at the institution attended.

Ethnicity or Race is defined as the ethnic characteristics the borrower self-identifies as being (U.S. Equal Employment Opportunity Commission, n.d.)

For-profit institutions are defined as an educational institution run by a profit-seeking business corporation (Council for Higher Education Accreditation, 2010).

Four-year institutions are defined as an institution that offers a bachelor's, master's, or another professional degree.

Generation "X" is defined as individuals born between 1966 through 1980 (Catalyst, 2012; Gelston, 2008; Theilfodlt & Scheef, 2004).

Graduate is defined as an individual who completed his or her program of study earning a certificate or diploma.

Less than two-year institutions are defined as an institution that offers a series of certificates that require less than two years of study.

Loan servicer is defined as the entity responsible for contacting and collecting payment(s) from the borrower on any federal loans the company holds or is contracted to collect.

Major is defined as the field of specialization or study in which a student is obtaining a certificate or degree.

Millennial generation is defined as individuals born between 1981 through 2002 (Catalyst, 2012; Gelston, 2008).

Non-graduate is defined as an individual who failed to complete his or her program of study.

Private institutions are defined as an educational institution funded and run by private individual(s), a religious organization, or by a nongovernmental agency (Common Data Set Initiative, n.d.).

Program of study is defined as the type of degree or certificate the student is pursuing.

Public institution is defined as an educational institution run by "publicly elected or appointed officials and derives financial support from public funds" (Department of Education, n.d., para. 5).

Stafford Loan (subsidized and unsubsidized) is defined as a federally guaranteed student loan under the Family Federal Education Loan Program or Direct Loan program with fixed interest rates.

Two- or three-year institutions are defined as an institution that offers a two- or three-year certificate or associate degree.

Type of institution is a variable under investigation and is defined as private, for-profit, or a public institution of higher education.

Theoretical Framework

There are numerous reasons why individuals default on their Stafford student loan debt. These reasons range from an inability to pay in a struggling economy, to lack of a financial support network, to beliefs that the education was not adequate, thereby justifying their refusal to pay back the obligation (Christman, 2000; Coleman, 2010; Flint, 1997; Jenkins & Lynch, 2007; Volkwein et al., 1998). Given the exploratory nature and scope of the study, there are many theoretical perspectives available to guide the researcher in understanding and explaining the factors that contribute to borrower default. As no two borrower's repayment experiences are alike, there is a need to present theoretical frameworks stemming from both sociological and economic contexts.

Economic

There are three theories derived from an economic basis—human capital theory, ability to pay, and debt burden. These theories help to explain why borrowers assume the financial burden and the reasons why a borrower might subsequently struggle with repaying the debt.

Human capital theory. In an economic sense, the term capital is widely reserved for a physical or financial investment that improves a business or individual economic outlook through increased output (Bottone & Sena, 2011). Becker (1994) posited that human capital is an intangible investment on a human level; such as education or training, that yields financial gains for both the individual and potential employers. Human capital differs from traditional economic capital in that the investment cannot be separated from the person (Becker & Mulligan, 1997). From this perspective, the investment and financing of an education are the rationalizations of a cost-benefit analysis of short and long-term financial goals of the individual (Lochner & Monge-Naranjo, 2004).

Ability to pay theory. The ability to pay theory has two components. The first component identifies a lack of financial resources as a reason to borrow funds to finance an education (Cabrera, Stampen, & Hansen, 1988; Cunningham & Kienzel, 2011; Flint, 1997). If the anticipated financial benefits outweigh the total costs of the program; the ability to pay theory states that, a person with limited financial resources will justify taking a loan(s) to achieve their educational goal. The second component of the ability to pay theory addresses the borrower's resources to repay the funds borrowed. The theory suggests that individuals with sufficient income or with financial support from friends and family are capable of paying back borrowed funds if the total resources are in excess of monthly repayment requirements (Christman, 2000; Volkwein et al., 1998).

Debt burden theory. The rates and total amounts an individual will borrow vary depending upon the unique circumstances that surround the person. Debt burden theory states that there is a threshold of 8% of monthly income between what are termed as manageable and unmanageable debt levels (American Council on Education, 2004; Baum & Schwartz, 2005; Cunningham & Kienzel, 2011; Kesterman, 2005). As college costs continue to increase, and more students fund their education through loans, the debt to income ratio is soaring.

Mishory and O'Sullivan, (2012) examined the student loan debt load of the graduating class of 2004 and found that the average borrower had a debt to income ratio of .49 when the prospect of a home mortgage was included. Most lending agencies require a debt to income ratio of 41% or less to qualify for a home, thereby, making the average 2004 college graduate ineligible to receive a home loan (Mishory & O'Sullivan, 2012). Increasing student loan debt levels can have lasting impacts on the financial wellbeing of a borrower, requiring those in repayment to delay or ultimately forgo the ideals of home ownership, car loans, and potentially starting a family. The delay in these goals caused by high debt loads can have effects on other sectors of the United States economy.

Sociological

In addition to economic factors, a number of sociological factors might also relate to why a borrower might default on the loan debts acquired through the course of schooling. These theories are student departure, approach-avoidance and the general model for assessing change.

Student departure theory. The theory of student departure analyzes the degree-of-fit between the student and their experiences at the institution (Flint, 1997; Tinto, 1987). To a degree, repayment behaviors portray the borrower's overall satisfaction with the quality of education and experiences acquired at the institution (Christman, 2000; Flint, 1997). This theory posits that students who believe they did not obtain a quality education or those who possess negative experiences of the institution are more likely to refuse repaying the loans received (Cabrera, Stampen, & Hansen, 1988; Coleman, 2010; Jenkins & Lynch, 2007). Whereas, those with positive experiences are more likely to repay the funds.

Approach-avoidance theory. The tenants of approach-avoidance theory state that how a student copes with stressors in an environment is directly related to the context and the individual's behaviors (Eaton & Bean, 1995). Lazarus, Averill and Opton (1974; as cited in Eaton & Bean), proposed that, depending upon the situation and context, the student's reaction to the stressor will lie on a continuum of approach (positive response) or avoidance (negative response) to the stimuli. In other words, a student will respond to a stressor either negatively or positively. As students enter college, the student will have either a positive (persistence) or negative (withdrawal) reaction to the environment. As research has shown, graduation is a strong predictor against student loan default (Gross et al., 2009).

Approach avoidance is also applicable to loan repayment. As students graduate and enter into repayment, the size of the monthly payment obligation, the timeframe to repay the amounts borrowed, or uncertainty on how they will repay on a limited income can trigger approach-avoidance behaviors. The larger the amount borrowed, number of lenders, or lack of awareness of repayment options can cause borrowers to either seek out resources (positive) or refuse (negative) repayment of their loan debts.

General model for assessing change. The students' developmental responses to the environment and organization's structural characteristics are the premise of the theory of general model for assessing change (Flint, 1997; Pascarella, 1985; Volkwein et al., 1998). This model assesses how well the student responds to the organization's characteristics such as mission, values, size, and structure (Pascarella, 1985). The more aligned and integrated the student is with the institution the more likely he or she is to repay the loans he or she received to finance his or her education. If the student borrower is unable to integrate him or herself into the institutional environment, it is more likely that the student will depart from the school in search for a different option. This departure, or student withdrawal, has been shown in previous studies to be a factor in student loan default (Gladieux & Perna, 2005; Kesterman, 2005; Volkwein et al.).

Determining Appropriate Theory(ies) for the Study

As demonstrated, there are both social and economic theories that help to explain why a borrower might subsequently default on his or her Stafford student loan debt. The relevant theory(ies) applicable depend largely upon the findings of this study. For example, if the resulting model shows that last reported enrollment status and total loan debt are key indicators of loan default, one could draw inferences to both economic and social theories. Applicable economic theories are debt burden and ability to pay. Applicable social theories are student departure and approach avoidance. If the resulting model shows that type of institution and last reported enrollment status are key indicators, the appropriate theories will stem from the social sphere, not the economic.

Delimitations

Delimitations are those factors that affect the generalizability of the findings that the researcher knowingly excludes from the scope of the research study. Delimitations of this study included a review of only those factors that yield a positive or a negative correlation to predicting student loan default.

Previous studies conducted on the topic have shown that certain factors of student demographic information play a contributory role in the repayment behaviors of the borrowers. Findings from the study conducted by Volkwein et al. (1998) showed that the student's ethnic background has minimal impact on default; but that age, the number of dependents a borrower might have, and gender have greater impact. Gross et al. (2009) noted that total amount borrowed, program of study, and whether the student graduated or not were predictors of default. Delimitations of this study consisted of reviewing defaulted borrower demographics including total undergraduate Stafford student loan debt, year of birth, gender, last reported enrollment status at that institution, and the total number of loan servicers the borrower has assigned to their loan portfolio.

By delimiting the study to reviewing only undergraduate student loan debt, graduate borrowers with typically higher loan debt were excluded from the study. Ethnicity and residency status were excluded from this study as variables. Ethnicity was excluded because of the variable being a voluntary self-reported field that students opt either in or out of providing at the time of submitting an application for admissions to the institution. The reliability of the data cannot be verified. Therefore, the variable was excluded from inclusion in this study. Residency was excluded because of the two-year institution's relaxed residency requirements in which a student must only demonstrate living within the area for six months to obtain in-county residency. This policy provided in-county tuition rates to students regardless of whether the student's intent was to return out-of-county or out-of-state after graduation.

Finally, the study was delimited to the findings of the data, and this research did not investigate the borrower's reasons for default or beliefs about the student loan program. If the study showed that a correlation among the variables exists in determining repayment status, future research can then be conducted to determine the factors in which students become delinquent, develop best practices for default management programs, and investigate the meaning of student loans to borrowers in repayment.

Limitations

Limitations of the study highlight restrictions of the research design and subsequent findings. The data were limited to the variables as identified in the research questions. The reliability of the findings is limited by the accuracy of institutional reporting to maintain the National Student Loan Data System (NSLDS) database. This limitation of the study was unavoidable; however, the researcher anticipated minimal inaccuracies due to the federal requirements of aid recipient reporting (Department of Education, 2011a; National Student Loan Database System, 2012).

A further limitation of this study was the use of the logistic regression statistical technique to identify highly correlated variables to a default or non-default status. The technique is sensitive to data outliers, missing data, and high correlations between variables resulting in multicollinearity and invalid results (Mertler & Vannatta, 2010).

As a result, if any of the variables showed a strong correlation to another variable, one of the variables would have needed to be removed from the analysis.

Due to the variables selected and unique nature of each borrower, normality was not expected in the data set. Logistic regression does not require normality within the predictor variables, and if normality was present, a discriminant analysis technique might yield a stronger analysis. An additional limitation of using logistic regression was that due to the combinations of discrete variables, there must be a large enough sample size used to limit standard errors and large parameter estimates (Mertler & Vannatta, 2010). As logistic regression analysis includes the goodness-of-fit test, the researcher needed to review closely the expected frequencies of each cell. Mertler and Vannatta recommend that "all pairs of discrete variables should be evaluated to ensure that all cells have expected frequencies greater than 1 and that no more than 20% have frequencies less than 5" (p. 293). Failure to assess the expected frequencies can result in the power of the analysis to be weakened.

Justification and Significance of Study

Between 2000 and 2011, there was a five-fold increase in total student loan debt (Goodwin, 2011) and an almost doubling of the national student loan default rate from 4.6% in 2005 to 10.0% for 2011 (Federal Student Aid, 2010b). Given that the total student loan debt amounts to a one trillion dollar industry, the total number of defaulted education debts is staggering.

Few studies have attempted to develop a reliable predictive model of student loan default over the last two decades (Flint, 1997; Fraas, 1989; General Accounting Office, 1991; National Summit of Councils, 2001; Volkwein et al., 1998). The applicability and validity of these findings are uncertain considering the substantial increase in student loan borrowers, total loan debt, and the differences in historical and societal contexts in which these studies were performed. Federal regulations state that if an institution's three-year CDR reaches a threshold of 30% the institution is at risk of losing its Title IV federal aid program participation agreement (Department of Education, 2011a). Facing the possibility of losing participation in federal aid programs, institutions need an updated model to assess and develop default management programs that adequately identify and target populations most likely to benefit from outreach opportunities.

The outcomes of this study are not intended to isolate or deny borrowers education loan opportunities based on borrower characteristics, but to aid stakeholders in identifying target populations for delivery of default management services. The findings of this research will help highlight problematic areas of federal policy and assist institutions, lenders, servicing agents and the Department of Education in accurately classifying at-risk populations, and develop best practices for default management delivery.

CHAPTER TWO: LITERATURE REVIEW

Popular media, consumed by high unemployment rates and federal budget crises, are fixated on the current state of the US economy (Chopra, 2012). Multiple programs administered under the Title IV federal regulation (34 CFR §668) include the Federal Work-Study program, Federal Supplemental Education Opportunity Grant, Federal Perkins Loan, and the more widely recognized Pell Grant and Direct Stafford Loan programs. The issue of student loan debt continues to gains momentum in national newscasts and the political arena due to increasing default rates, high unemployment adding to the inability of borrowers to repay, and the economic ramifications of Title IV federal financial aid program integrity and funding streams.

As a result of these concerns, regulatory changes increase the accountability and impose stricter sanctions on higher education institutions to ensure current students and student loan borrowers comply with program rules over the lifespan of their student loans. This section discusses gaps in the available research, how this study adds to the knowledge base, a review of the literature regarding national, institutional, and state-centric default characteristics, and recent regulatory changes affecting Title IV programs. Following those sections is a discussion on the importance of the study to develop a model of student loan default characteristics, what a model is, and finally the methods to carry out the study.

Gaps in Available Research

Recent research on student loan default models has focused largely on institutional and state-centric default borrower models (Gross et al., 2009). The reason for this transition has been the loss of national databases that warehouse both institutional and borrower-based demographics. Relevant research on a national scale is dated with the majority of studies stemming from the early to mid-1990s (Flint, 1997; General Accounting Office, 1991; Volkwein, Cabrera, Szelest, & Napierski, 1998).

Another reason for the loss of relevant research on a national scale could be the result of complex and changing federal regulations holding institutions of higher education more accountable for their CDRs. That being said, institutions are more interested in understanding the demographics of their own borrower population and not as willing to invest time and resources in understanding the national loan portfolio. Specific information regarding their borrower characteristics is viewed as more useful to institutions when developing and implementing default management operations.

Increasing student loan default rates, a struggling economy, and changing federal regulations highlight a need for additional research on student loan borrowers. Research is needed on a widespread scale including, but not limited to, default borrower characteristics, financial literacy programs in higher education, borrower repayment experiences, and investigating the reasons why borrowers default on student loan debt. The scope of this study is to add to the knowledge base information on default borrower characteristics.

How this Research Adds to the Knowledge Base

Title IV federal financial aid programs administered by the Department of Education (ED) continue to be threatened because of the current economic crisis, sequestration, and widespread budget cuts. This research helps fill a void in the available knowledge about the present day characteristics of defaulted student loan borrowers. The outcomes of this study add to the knowledge base by:

- adding the millennial generation into the equation of borrower repayment trends;
- providing an updated model in understanding what characteristics a defaulted borrower has in the institutional student loan portfolio;
- providing a framework for institutions of higher education, government agencies, and loan guarantors to determine default borrower characteristics for their loan portfolios;
- serving as a statistical baseline to establish effective financial literacy and default management programs to reduce CDRs; and
- assisting government and regulatory bodies in evaluating the impact of Title IV regulations, evaluating student borrowing practices, and critically reviewing current federal policies.

Default Characteristics

The reasons for delinquency and subsequent default vary depending upon the borrower; however, there is some evidence that characteristic relationships exist. The variables under investigation for this study have been identified from previous research as having strong correlations to default behavior. This section is structured first to introduce default borrower demographics, then institutional influences, socioeconomic factors, and state-specific research into default characteristics.

Borrower Demographics

Prior studies show relationships between default and individual demographics such as ethnicity, age and gender (Barone et al., 2005; Gross et al., 2009; Kypuros, 2009; Lee, 2009).

Ethnicity. In a number of studies, African Americans were found to have the greatest risk of default (Harrast, 2004; Herr & Burt, 2005; Steiner & Teszler, 2003). Volkwein et al. (1998) also noted that African Americans were less likely to rehabilitate their student loans once they entered into a default status. Reasons for why this phenomenon appears to hold true are unknown; however, Boyd (1997) postulated that interactions with market discrimination in the mortgage industry facilitate a loss of desire to protect ones' financial credit score resulting in subsequent default on student loan debt. Volkwein et al. noted that African Americans tended to borrow more heavily during periods of personal, family, or employment hardships than their counterparts did and that upon departure they tend to be more likely unemployed. Volkwein et al. attributed a diminished capacity to repay the loans due to limited employment opportunities and less satisfaction with their education experience as possible reasons for default.

Age. In a review of the literature on student loan default characteristics, Gross et al. (2009) found that nearly all studies under investigation noted that as age increases the likelihood of default also increases. Even when controlling for the variable of income, findings show that as age increases so does the likelihood of default (Christman, 2000; Flint, 1997; Herr & Burt, 2005; Steiner & Teszler, 2003). Herr and Burt postulated that the reason for this trend was a result of having more financial responsibilities—such as having families, which younger borrowers are less likely to have started. This is important to note, as household size and the number of dependents a borrower reports has been shown to be a predictive variable in loan default (Gross et al., 2009). Harrast (2004) and Choy and Li (2006) credited an increased overall debt burden as the reason why older borrowers encountered more difficulties in student loan repayments.

In opposition to older studies, Cunningham and Kienzel (2011) found that for borrowers who entered repayment in 2005, a younger generation—the millennials, were more likely to become delinquent or default on their student loan debt. Data from this study showed that almost half of the population under review was between the ages of 21-30 and that older borrowers did not experience similar problems with repayment behaviors (Cunningham & Kienzel, 2011). This shift in repayment behaviors between the generations could be the result of a lack of financial literacy, limited job market skills, and unfamiliarity with repayment options, which an older borrower would have determined over time (Cunningham & Kienzel, 2011).

Gender. The relevant literature on gender points to a less distinct correlation with default. Some studies show that men are more likely to default than women are (Flint, 1997; Podgursky et al., 2002; Woo, 2002). Other studies highlight a lack of difference in default patterns between the genders (Harrast, 2004; Volkwein & Szelest, 1995). When controlling for income, Shwartz and Finnie (2002) found no significant difference between the genders and default patterns. However, Choy and Li (2006) did find that women on average take longer to repay their student loan debt when compared to their male counterparts.

Institutional Characteristics

Research on institution characteristics shows that type of institution attended, choice of major/program of study, and degree completion strongly correlate to a borrower's likelihood to default on their Stafford loan debt(s). Other factors such as college grade point average, tuition, and fees have been shown to have a correlation with student success and borrower repayment.

Type of institution. Recent studies show that the type of institution last attended is a predictor of future loan default (Cunningham & Kienzel, 2011; Gladieux & Perna, 2005; Kesterman, 2005). In those studies, results showed that students who attended for-profit or two-year public institutions had higher instances of default than their four-year public institution counterparts did. These findings seem to hold true as older studies came to the same conclusion that less-than-two-year, proprietary, and two-year community college borrowers were most likely to default on their student loan debt (Podgursky et al., 2002; Woo, 2002). Other studies that took into account institutional and demographic characteristics, borrowing behaviors, and income found that the differences in type of institution attended and default largely disappeared (Flint, 1997; Volkwein et al., 1998).

Major selection and program of study. In a study conducted by Mullins (2008), findings showed that undergraduate students closely tied their total loan debt to their program of study and projected income. She posited that students with an expected higher projected income borrow more loans than those students whose projected income would be lower than their peers' would (Mullins, 2008).

In a study looking at Texas A&M loan defaulters, findings suggest that a student's choice of major is significantly tied to the likelihood of loan default (Steiner & Teszler, 2003). From the findings, this researcher noted that the more specific the major, the less probable the student will subsequently default on his or her loans. Students who selected a major in marketing or accounting had respective loan default rates of 1.6% and 1.8% (Steiner & Teszler, 2003). In comparison, for students who selected a more generic major, such as general studies or business administration, their likelihood to default increased to 14.7% and 13.6% respectively (Steiner & Teszler, 2003). An anomaly to their study was students who chose an interdisciplinary major with a 1.9% chance of default. This researcher speculates that the low probability of student loan default for interdisciplinary studies majors is due to the specific and tailored nature of the program. The focused and narrow program design could allow the borrower to develop a market niche upon graduation, thereby, increasing the likelihood of loan repayment; however, further studies would need to be conducted to accept or reject this hypothesis.

Program completion. Research suggests a negative correlation between default and successful completion of the program (Christman, 2000; Cunningham & Kienzel, 2011; Kesterman, 2005; Loonin & McLaughlin, 2012). Christman (2000) noted that students who remain enrolled for more than two semesters were more likely to avoid default than those who withdrew within their first year. Failure to remain enrolled at the institution could be indicative of the student's and institution's failure to adequately integrate within the culture of the institution or a financial or personal hardship.

Studies conducted by Volkwein and Szelest (1995), Gladieux and Perna (2005), and Steiner and Teszler (2003) found a positive correlation between default and students who withdrew and failed to attain a certificate or degree. These findings suggest that borrowers who complete a degree or certificate are least likely to default. These researchers further postulate that when a borrower earns an advanced degree, the borrowers are more likely to maintain the monthly payments obligations as a result in higher income potential, preventing loan default.

College GPA. College GPA has shown to be correlated with program completion. Research shows that the higher the GPA, the less likely the borrower is to default on student loan debt (Flint, 1997; Volkwein et al., 1998). Additional studies show similar findings in that the higher the GPA, the more likely the borrower is to complete his or her program of study and less likely he or she is to default on his or her educational debt (Gross et al., 2009). In a logistic regression study looking at default characteristics at a two-year community college, findings showed that if the borrower's cumulative GPA is less than 2.00 there was a 55% chance of student loan default within a three-year timeframe (Christman, 2000). Not surprisingly, if the student failed any classes, there was a 69.2% chance of default (Christman, 2000).

In a study conducted by Kantrowitz (2009), findings showed that borrowers over the age of 24, with a less than 3.00 GPA, tuition and fees greater than $5,000, and low income were more likely to incur student loan debt of $40,000 or more. While research was not conducted to investigate whether these borrowers subsequently defaulted on their educational debt obligations, the ability to pay theory suggests that these borrowers might struggle with loan repayment due to low income and limited financial resources.

Tuition and fees. In a study examining factors leading to excessive student loan debt from borrowers in the 2003-2004 and 2007-2008 National Postsecondary Student Aid Study, tuition and fees greater than $5,000 was identified as one of the reasons for increased borrowing (Kantrowitz, 2009). Kantrowitz determined that for "every $2,500 increase in the cost of attendance translates into a $725 increase in annual debt" (2009, p. 8). As tuition and fees increase, so do the economic resources required for borrowers to cover their cost of education.

Federal financial and institutional grant aid resources have a strong correlation with reduced overall loan debt and college persistence toward a degree (Coleman, 2010).

Christman (2000) identified that if borrowers lose federal student aid eligibility, the likelihood of default increases by 39.2%. Investigating the financial ability of students attending private education institutions, Coleman (2010) found that institutional tuition reduction programs increased the likelihood of long-term student enrollment and persistence to a degree. As previously identified, degree completion is correlated negatively with default (Christman, 2000; Cunningham & Kienzel, 2011; Kesterman, 2005; Loonin & McLaughlin, 2012).

Socioeconomic Factors

In addition to borrower demographics and institutional characteristics, a host of socioeconomic factors also shows correlations to student loan default statuses. Research highlights variables such as income and debt burden as potential contributors to a default repayment status.

Income. Students from higher family income typically have a lower probability of defaulting on their student loan debt upon entering repayment (Woo, 2002). This is likely due to additional household resources available to the borrower during times of personal financial hardship. Students from low-income households have tendencies to borrow more during school (Herr & Burt, 2005; Steiner & Teszler; 2003). Not surprisingly, low-income students, with a higher debt load, report feeling financially burdened upon entering repayment (Baum & O'Malley, 2003; Gross et al., 2009). Choy and Li (2006) and Lochner and Monge-Naranjo (2003) noted that as post departure earnings increase the likelihood of default decreases over time. Feelings of being financially burdened upon entering repayment could be a result of low-earning entry-level positions or a lack of a financial safety net during personal hardships (Woo, 2002).

Debt burden. In general, a high total student loan debt when compared to earnings has been shown to be strongly correlated with student loan default (Schwartz & Finnie, 2002). In a study conducted by the California Postsecondary Education Commission (2006) looking at the total debt of burden of students in the 2003-2004 aid year, findings showed that borrowers who attended for-profit institutions owed over $38,000 at time of repayment. This is in contrast to students who attended private four-year institutions owing on average $36,000 (California Postsecondary Education Commission, 2006).

Interestingly, Cunningham and Kienzel (2011) found that, on average, borrowers who entered repayment in 2005 with three or fewer loans for a combined total of $8,000 or less were more likely to default than borrowers with higher debt and a greater total number of loans. These findings were also similar to those of Steiner and Teszler (2003) and Woo (2002). Cunningham and Kienzel postulated that the number of years the borrower was enrolled, program completion, and type of institution last attended were underlying factors that might have facilitated a default repayment status. As previously mentioned, both program completion and types of institution last attended are strongly correlated with a default repayment status.

Another exception to this finding is if the high loan debt is because of pursuing a graduate program (Volkwein et al., 1998; and Woo, 2002). Borrowers who achieve a graduate degree or higher are usually able to earn a higher income than their counterparts who failed to achieve a degree or achieved a certificate, associate, or bachelor's degree.

State-Specific Default Borrower Research

A relatively new area of research includes the state-centric approach to investigating student loan repayment issues. Findings in this area of research are meant to highlight statewide trends in repayment status and evaluate the effectiveness of policies and programs.

Nevada. With historically high statewide default rates, a need was present to investigate the state's default borrower population characteristics. A regression analysis was conducted to investigate relationships between institutional and borrower demographic characteristics to determine whether a model could be developed to determine likely candidates of future default (Kypuros, 2009). The dependent variable was loan repayment status while the independent variables under investigation were age, gender, ethnicity, graduation rate, residency, and major.

Findings from a simple linear regression analysis showed that age and residency status strongly correlate with default rates at the statewide level (Kypuros, 2009). The variable of age was shown to be related negatively to default rates while residency status was correlated positively to default rates. Age, graduation rates, and ethnicity were shown to have a strong association with default at the institutional level (Kypuros, 2009). At the regional level, multiple regression analysis shows strong associations among age, social science and science majors, to default rates (Kypuros, 2009). The findings of this analysis, with regard to age, are in contrast with previous findings that show a positive relationship to default as age increases. Kypuros' (2009) findings on graduation rates having a positive association with reducing default further supports previous research with similar findings (Gladieux & Perna, 2005; Steiner & Teszler, 2003; Volkwein & Szelest, 1995).

Florida. In an evaluation of default management practices among Florida's two-year community colleges, Salas-Amaro (2008) determined that there was an overall lack of standardization in policies and practices among the 20 institutions that were interviewed. The majority of institutions interviewed completed the bare minimum federal requirements, which require borrowers to complete loan entrance counseling before receiving a loan, and to encourage students to complete loan exit counseling upon graduating or withdrawing from the institution.

Of interest was that multiple two-year Florida community colleges send additional mail outs to students who drop below half-time enrollment (Salas-Amaro, 2008). The reasons why financial aid offices struggle with or refuse to implement more rigorous default management policies and practices were beyond the scope of this study. Based on the study findings, Salas-Amaro proposed an eight-step default management program to facilitate the standardization of default management practices among Florida's two-year community colleges. The proposed plan is presented below:

1. entrance counseling;
2. continued student loan enrollment;
3. repayment of student loan workshops;
4. classroom visits and faculty involvement;
5. exit counseling;
6. skip tracing;
7. ad-hoc reporting; and

8. National Student Loan Database System report (Salas-Amaro, 20078).

The plan encompasses the minimum federal requirements for entrance and exit counseling, but also strengthens the default management program (Salas-Amaro, 2008). The proposed model includes ongoing informational sessions while borrowers remain in school and includes three different reporting measures to notify the Department of Education, lenders, loan servicers, and guarantor agencies of enrollment status and changes to borrower contact information.

California. California boasts the largest community college network within the United States. Within this network of two-year colleges, institutional default rates vary widely with some low while others are high. In a landmark study comparing policies and procedures among community colleges with similar student borrower demographics, Smith (2002) concluded that there were slight statistically significant differences in institutional policies, procedures, and practices that could account for the wide difference in default rates. Institutions employing the bare minimum federal guidelines had higher loan default rates; whereas institutions employing both institutional and federal procedures had lower default rates (Smith, 2002).

Of particular note is that Smith (2002) concluded that institutions with lower default rates had more available resources in terms of staffing, monetary funding, and college support as opposed to those institutions with higher default rates. While the findings from this study were not determined to be significant in accounting for the fluctuation in default rates among similar institutions, this does lead one to question the larger impacts that staffing models, monetary resources, and college support have on department operations. One possible reason for these differences could be that financial aid personnel and other college staff were more readily available to assist borrowers with questions, provide information on loans, and offer feedback on borrowing habits.

Regulatory Changes

Amid economic instability and high unemployment rates, the President and Congress have been tasked to reduce government-held debt. Over the last couple of years, regulatory changes have greatly increased the administrative burden on higher education institutions, lenders, guarantee agencies, and the US government officials who administer Title IV federal financial aid. The majority of changes have been designed to tighten financial controls over the various Title IV aid programs.

Budget Control Act

The Budget Control Act (2011) eliminated both the graduate subsidized loan program (§502), and direct loan repayment incentives for on-time loan payments (§503) at the onset of the 2012-2013 fiscal year. In addition, the Budget Control Act (2011) established the formation of a congressional super committee tasked to develop a plan by November 2011 to reduce the federal budget deficit by 1.2 trillion dollars or additional cuts across the board would take effect for the 2013 fiscal year. The inability of the congressional super committee to develop a deficit reduction plan directly affects future funding for the Federal Supplemental Educational Opportunity Grant and Federal Work Study programs (National Association of Student Financial Aid Administrators, 2011). As the federal budget crisis worsens, further regulatory changes to reduce expenses are expected within higher education financial assistance programs.

Program Integrity

Access to higher education and the continued funding of Title IV programs to help pay for the cost of education are areas of increasing interest for post-secondary institutions, students, and the public at large. Among the regulatory changes higher education financial aid administrators face are to strengthen program integrity—including implementing gainful employment requirements (Program Integrity Issues, 2010), and address the multiple issues resulting from changes to the direct loan programs (Cohort Default Rates, 2009; and Two Year Cohort Default Rates, 2000). Designed to strengthen the Title IV aid programs, these changes help prevent fraud and abuse of the system, encourage students to make more educated and informed decisions about their borrowing habits and selection of a degree program, while holding institutions accountable for continued program participation.

Amendments to the Higher Education Reauthorization Act

Substantial changes to the direct loan program result from the October 2009 amendments to the Higher Education Reauthorization Act of 2005. As part of the Higher Education Act of 1965, higher education institutions are held responsible for their former student's CDR and could suffer steep sanctions such as loss of program participation eligibility if the CDR exceeded 20% (Cohort Default Rates, 2009; Two Year Cohort Default Rates, 2000). The 2009 amendments change the CDR calculation from a two-year to a three-year rate and establish new threshold guidelines for applying institutional sanctions (Cohort Default Rates, 2009; Two Year Cohort Default Rates, 2000). The new regulations state that if the default rate exceeds 30% for three consecutive years, the institution will lose the opportunity to participate in Title IV federal financial aid programs (Pierson, Walsh, & Wright, 2011).

Direct Loan and Federal Family Education Loan Program (FFELP)

Previously, institutions offering Stafford and PLUS education loans were able to participate in either the FFELP or Direct Loans programs. Both programs offered the same repayment terms for student loan borrowers. The Direct Loans program allows students to borrow directly from the US government, whereas FFELP provided students with the option to choose independent lenders to administer and service their US government guaranteed Stafford loans (Department of Education, 2010).

Before the economic crisis of 2008, lenders under FFELP would compete for both student borrower and institutional program participation by offering an assortment of repayment incentives for on-time repayments as well as providing delinquency and default management services for borrowers. In 2010, the US government eliminated the Stafford and PLUS loan borrowing options under FFELP (Student Aid and Fiscal Responsibilities Act, 2010). Individuals looking to finance their education through Title IV federal aid programs now borrow directly from the US government under the Direct Loan program. This regulatory change resulted in the loss of repayment incentives for on-time payments and shifted delinquent and default management operations directly onto the institution of higher education offering Direct Loans to students to assist in covering the cost of education.

Consequences of High Cohort Default Rates

Increasing default rates and the loss of eligibility to participate in Title IV federal aid programs can have immediate and direct ramifications for a number of stakeholders.

Students who rely on Title IV aid to pay for the costs associated with earning a degree might be unable to afford their program financially and subsequently drop out without earning a degree.

Pell

Revenue funds from the Stafford Loan Program are used to help maintain and safeguard the funding needs of the Pell Grant (New America Foundation, 2013). As delinquency and default rates increase, Pell funding increases to offset inflation might have to constrict to reflect available balances adding increased financial hardships on low and middle-income students (New American Foundation, 2013).

Interest Rates

Backed by the full faith of the US government, the Stafford loan program is funded through US Treasury bills (Higher Education Act, 1965). As default rates increase, national interest rates might also increase to account for the loss in expected revenues creating an additional burden on US taxpayers.

Institutional Consequences

The Higher Education Opportunity Act of 2008 resulted in the calculation of a three-year cohort rate and the establishment of a 30% threshold for sanctions (Federal Student Aid, n.d.). These changes took effect in 2012 based on loans that entered into repayment in 2009 (Federal Student Aid, 2010a). In 2010, the Department of Education conducted a test run of 2007 figures to illustrate the change these amendments would have to the national and institutional CDR values (Pierson et al., 2011). Where previously only 5 schools received sanctions for exceeding the former 2-year 20% CDR regulation, the test run of 2007 figures showed 315 institutions will have an expected default rate at or above 30% with another 808 institutions possessing a 20% to 29.9% projected CDR (Federal Student Aid, 2010a). Current reports show that over 300 institutions are in jeopardy of losing title IV program participation, as their three-year CDRs have exceeded 30% every year for the last three years (Federal Student Aid, 2014).

Loss of program participation eligibility could affect an institution's financial ability to continue operations. A decrease in enrollment resulting from students unable to pay out-of-pocket for tuition and fees might result in fewer program offerings and the subsequent closure of institutions that are unable to adapt and compete in a new environment.

Local, State, and National Economic Effects

As the economy continues to remain weak, coupled with a 9% national unemployment rate (Department of Labor, 2011), more and more borrowers are becoming delinquent and subsequently defaulting on their Stafford Education Loans. In a study investigating the prevalence of delinquency and default repayment statuses, for every one borrower in default there are two borrowers delinquent (Cunningham & Kienzel, 2011). As a result of the changes to the CDR calculation, many are predicting a substantial increase in institutional cohort default values, placing a number of institutions in jeopardy of losing their participation agreements with the Department of Education by 2014 (Federal Student Aid, 2014; Lederman, 2011; Office of Scholarships and Financial Aid, 2011).

Depending upon the size of the institution(s) and the health of the local economy, the loss of millions of dollars in semi-annual aid disbursements could greatly diminish constituent purchasing power further harming the local community. On a widespread scale, depending upon the number of institutions and prevalence of the problem, both the state and national economies could feel these consequences. As stated, the Direct Loan program is backed by the full faith of the US government and is funded through annual auction bids of the US Treasury Bill. As Direct Loan default rates continue to rise, both domestic and foreign investors alike could view annual auction bids for the US Treasury Bill as a risky financial venture and invest funds in other fiscally viable opportunities. Failure to attract investors for US Treasury bills could result in an increase in national interest rates creating the climate for inflation and stagnation as the economy remains weak and unemployment remains high.

Importance of Study

A need is present to understand better the makeup of an institution's student loan portfolio, borrower characteristics, and to explore the development of a predictive model of student loan default. Findings from this research help highlight baseline borrower trends in the student loan portfolio specific to age, gender, graduation rates, total borrowed debt while at the institution, and last reported enrollment status at the institutional level. The research also highlights a much needed knowledge area for state and federal government policymakers looking to evaluate the effects of Title IV regulations, understand student borrowing practices, and critically reviewing current state and federal educational policies intended to promote the overall academic and financial wellbeing of students. Furthermore, any findings could be used to support or reject previous research on borrower demographics and institutional characteristics in relation to student loan default.

The development of a predictive model of student loan default would aid a number of stakeholders looking to benchmark results against, or develop, other student loan default models. Stakeholders who might be interested in the development of a model include individual institutions, sector-specific groups such as guarantee agencies, statewide special interest entities such as two-year community colleges, for-profit or private education associations, state Department of Education offices, local and state legislators, and the federal government. To understand the value added when developing a predictive model, a brief discussion is needed about how to develop a model and assess its merit.

What is a Model

The purpose of research is to explore a concept, issue, or phenomenon. In some situations, research can be undertaken to solve a problem or to justify searching for outside assistance. All forms of research are driven by theory and based on a philosophical stance (Shafique & Mahmood, 2010). By establishing research within a specific theory, it sets the stage for a common and agreed upon view of the phenomenon under investigation, explains why there is a need for further research, and sets the tone for how the phenomenon will be studied. One form of research that helps frame the issue under review is the development of models.

The purpose of a model is to provide the basic concepts and contributing factors that describe the phenomenon and the conditions under which researchers can study it (Clarke; 2005). As such, models are often presented as physical, verbal, graphical, or mathematical constructs representing a phenomenon. According to Stockburger (1996), all models are incomplete, as they do not include every aspect or possible outcome that exists in reality. While models are designed to facilitate understanding of a phenomenon or issue, they are simply a representation of reality and need to remain flexible to account for new information as it becomes available.

Assessing the Merit of a Model

All models have common core characteristics. Leimkuhler (1972) provided five characteristics of well-designed models. These characteristics include:

- how well the model relates to other available models;
- transparency and ease in understanding or interpreting the resulting model;
- ability to deduce information from the model;
- robustness or sensitivity to assumptions that can be derived from the model; and
- how easily the model can be modified and expanded based on new information (Leimkuhler, 1972; as cited in Shafique & Mahmood, 2010).

If the resulting model is missing one or more of the characteristics, the ability to derive information and expand research on the phenomenon is hindered. A model that is too narrow limits the applicability to expand further research in the area due to the constrictions of the model's design. Alternatively, a model that is too broad in scope fails to adequately provide a structured and agreed upon framework to view the phenomenon so further research can be conducted. In either situation, the resulting model could lead to overgeneralizations about the concept or issue being investigated. A model that meets the characteristics listed above offers a practical guideline that can be validated to show the interrelationships of actions and reactions within a phenomenon and remains flexible to new findings (Shafique & Mahmood, 2010).

Appropriateness of Methods

The relevant literature shows that the reasons for delinquency and subsequent default result from a highly complex system of borrower-based beliefs coupled with borrower- and institution-specific characteristics. This research did not attempt to identify the borrower-based philosophies about why a borrower defaulted but what the overall characteristic trends show that put a borrower most likely in jeopardy of defaulting on their loan debt. Instead, this research attempted to develop a predictive model of student loan repayment (default or non-default) from knowledge of specific variables. These variables include total undergraduate Stafford student loan debt, total number of Stafford loan services, year of birth, gender, type of institution attending in which loan was received, and last reported enrollment status at that institution. To select the appropriate research method the researcher needs to understand the intended outcome of the research along with the variables under investigation.

Research Design

The structural foundation that lays down the framework for how a study will be carried out is the research design. The research question(s) and study intent ultimately drive the research design of either a qualitative, quantitative, or mixed methods approach (Creswell, 2009). Qualitative research concentrates on understanding the meaning of a phenomenon and data collection occurs through observations, case studies, interviews, and surveys (Creswell). Quantitative research emphasizes the use of statistical tests to analyze variables to determine whether patterns and relationships exist within data (Mertler & Vannatta, 2010). A mixed methods approach uses both qualitative and quantitative approaches to search for relationships within data and to discover the prescribed meaning behind the phenomenon.

This study lent itself to a quantitative research design because the research questions are written to look at a multivariate analysis of what, if any, relationships exist between a number of variables. The dependent variable was a dichotomous categorical variable of default or non-default, and relevant literature of previous studies shows that the logistic regression analysis is a widely accepted research technique to compare variables to produce a predictive model of default repayment behaviors (Barone et al., 2005; Christman, 2000; Flint, 1997; Volkwein et al., 1998; Volkwein & Szelest, 1995).

Logistic Regression Outcomes in Relation to Research Questions

The output of each logistic regression test includes three main components. These components include a goodness-of-fit test, a classification table, and a table of coefficients. These three outputs allow the researcher to assess whether the variables included in the model accurately classify individual cases, the overall weight each variable has on the dependent variable, and how valid the overall model is at predicting default repayment status. Coupled with the bivariate analyses for each of the four groups (i.e., overall institutional, baby-boomer, generation X, and millennial generations) a more detailed profile of the defaulted student loan borrower is inferred.

Data Measurements

The NSLDS is a repository of information maintained by the United States Department of Education with borrower financial aid information updated no more than every 30 days as appropriate (Department of Education, 2011b). The information is provided to NSLDS by loan guaranty agencies, ED loan servicers, ED debt collection services, Direct Loan servicing, Common Origination and Disbursement, Conditional Disability Discharge Tracking System, Central Processing System, and schools (National Student Loan Database System, 2012). The NSLDS is unique in that it provides a historical and current record of a student's enrollment, financial aid history including grants and loans, loan balances and statuses, lender and guarantee contact information, student legal name changes, dates of birth, course of study, gender, institutions attended by the borrower, and more (National Student Loan Database System, 2012). Each institution of higher education is provided a three-year CDR report on an annual basis. The three-year CDR report contains the type, repayment status, and amount of each loan that was awarded to the borrower by the institution, last reported enrollment, loan servicer, lender code, and date of birth, among other variables. See Appendix A for a list of variables provided to the institution on the three-year CDR.

Default Rate Population

An institutional review board request was submitted to the host institution requesting access to their 2011 three-year CDR report from NSLDS. The host institution was a large two-year public community college located on the East Coast of the United States. Approximate enrollment size for this institution is 70,000 students annually. The 2011 cohort default population consists of borrower repayment activities over a three-year timeframe ending in 2013. This final number is used as a measure to assess the national default rate as well as to determine whether institutional sanctions are required for institutions that have exceeded the 30% CDR threshold.

Stratified Random Sampling

Through stratified random sampling, data obtained from the host institution were reduced to a manageable sample size that maintains a representation of the overall student loan borrowing population. Stratified random sampling is a technique that takes into account the probabilities of identified groups within the population to keep samples representative of the overall data set (Lund Research, 2012).

Mertler and Vannatta (2010) warned that when the ratio of cases to variables is small, large parameter estimates and standard errors are more likely. These types of errors ultimately result in the combinations of discrete variables and model bias. Model bias occurs when the resulting model is overfitted to the sample group resulting in inaccurate findings. Research suggests that to limit model bias a minimum of 10 to 20 events per variable is recommended (Courvoisier et al., 2010; Peng, Lee, & Ingersoll, 2002). At 6 total variables and 20 events per variable, the researcher needed a minimum of 120 cases for each logistic regression analysis to limit model bias.

CHAPTER THREE: METHODOLOGY

The quantitative research supported a logistic regression design because the purpose of the study was to analyze what, if any, relationships exist between several variables. Statistically significant findings would aid in the development of a predictive model of student loan default for borrowers who attended a two-year public institution. The forward logistic regression technique was selected due to simplicity in exploring the relationship(s) the variable(s) have in combination with the outcome.

Research Questions

1. Can a default loan repayment status be correctly predicted from knowledge of total undergraduate Stafford student loan debt, total number of Stafford loans servicers, year of birth, gender, and last reported enrollment status at that institution?
2. If a default loan repayment status can be predicted correctly, which independent variables are central to the prediction of that status?
3. How many defaulted borrowers are classified correctly?

Null Hypothesis

Loan repayment status (default or non-default) cannot be correctly predicted for an institutional model from knowledge of total undergraduate Stafford student loan debt, total number of Stafford loan servicers, year of birth, gender, and last reported enrollment status at that institution.

Research Design

The dependent variable under investigation was loan repayment status, which is a categorical and dichotomous variable with an outcome of default or non-default. The independent variables for this study were a combination of categorical, ratio, and interval variables. The independent categorical variables for this study were gender and last reported enrollment status at that institution. The independent ratio variable was the borrower's total Stafford student loan debt. The independent interval variables were the total number of Stafford loan servicers and year of birth.

Interval and ratio variables were already in a numerical format and did not require transformation. Understanding which variables were categorical was critical because they required numerical reclassification to be able to run the logistic regression analysis. Standardization of categorical variables is discussed later in this chapter.

Logistic regression models are not sensitive to nonlinear relationships between one or more independent variables, which add to the method's overall flexibility. To enhance findings, the researcher investigated a series of bivariate analyses based on age groups to identify means, frequency distributions, and standard deviations of the relevant variables for use in drawing inferences on what the average borrower and defaulted borrower looks like for each of the three generations.

Instrumentation

A variety of software systems exists to complete the data analysis component of this study. For the purpose of this research, SPSS 18.0 GradPack statistical software was used for all data analysis activities.

Population

The population under investigation was undergraduate Stafford loan borrowers, who attended a large two-year public community college on the East Coast and entered repayment in 2011. The loan status history was collected for each borrower over a three-year period. The three-year review mirrors current federal practice for analyzing both the national and institutional three-year CDR calculation. With institutional review board approval from the host and parent institution, a report was obtained through the school from the NSLDS to identify the population group under investigation.

The official 2011 Three-Year Cohort Default Report for the institution provided a number of variables under investigation. The report contained the borrower's name, social security number, year of birth, total undergraduate student loan debt, loan repayment status, last reported enrollment status, and number of loan servicers (see Appendix A for a full listing of variables). The two-year public institution's Planning, Research and Evaluation office cross-referenced the social security numbers with institutional records to provide borrower gender data, then removed borrower name and social security values and reassigned a random student identification number. Removing borrower name and social security number removed the chance for borrower identification and safeguarded borrower rights.

Sampling

Due to the size of the data set, stratified random sampling was used to keep the number of cases under review manageable while also keeping the data representative of the overall population. The relevant variables for the stratification process were gender and loan repayment status. Because the number of cases within the population was unknown, the researcher used a sample size statistical calculation to select an appropriate sample size that maintained a 95% level of confidence.

The first sample served as a learner group to develop the predictive model(s) as outlined in the research question. The second sample served as the test group to test the reliability and validity of the "learner" model against a new sample group of individuals in repayment.

Standardization of Variables

To conduct the analyses, each of the categorical variables, dependent and independent, were recoded in a numerical format. The categorical variables in this study included the dependent variable of borrower repayment status and the independent variables of gender and last reported enrollment status. Table 1 identifies each of the variable outcomes and their corresponding numerical assignment.

Table 1

Numerical Values for Variable Outcomes

Variable	Measurement	Coding
Loan repayment status	Categorical	0 = Non-default, 1 = Default
Gender	Categorical	0 = Male, 1 = Female
Enrollment status	Categorical	0 = Graduated, 1 = Withdrawn, 2 = Other
Year of birth	Interval	Open-ended 1900-2013
Total loan debt	Ratio	Open-ended 1-57,500
Number of loan services	Interval	Open-ended 1-100

Each of the categorical outcomes was assigned a numerical value to facilitate analysis. As an example, the three classifications for last reported enrollment status are graduated, withdrawn, and other. Graduated was coded as the default value of "0," withdrawn was coded with a value of "1," and other was coded with a value of "2" to differentiate the three categories from one another.

Descriptive Analyses Procedures

Prior to conducting the logistic regression analyses, a series of bivariate tests to obtain baseline information on the variables and trends within the data set was conducted. For the overall institutional model, the entire population was run for bivariate analysis looking at the mean, standard deviation, and frequency distribution of each variable against both default and non-default loan repayment statuses. The population data set was broken down into three different data subsets, determined by year of birth, to obtain the generational-specific trends. Any borrowers with a year of birth outside of the generations under review, born between 1946 through 2000, were not included in the study. Each of the generational population data subsets was run for bivariate analysis looking at the mean, standard deviation, and frequency distribution of each variable against both default and non-default loan repayment statuses.

Multivariate Analyses Procedures

Two forward logistic regression analyses were conducted to determine whether student repayment status (default versus non-default) could be predicted accurately from the independent variables identified. The first logistic regression analysis was to develop the predictive model of student loan default. The second logistic regression analysis tested the previous model's reliability and validity to ensure model accuracy on a separate sample group.

CHAPTER FOUR: RESULTS

This research explored whether a predictive model of student loan default could be developed from knowledge of five independent variables. To complete this study, a host institution provided access to their 2011 three-year CDR report. Research questions under investigation included:

1. Can a default loan repayment status be correctly predicted from knowledge of total undergraduate Stafford student loan debt, total number of Stafford loan servicers, year of birth, gender, and last reported enrollment status at that institution?
2. If a default loan repayment status can be predicted correctly, which independent variables are central to the prediction of that status?
3. How many defaulted borrowers are classified correctly?

Logistic regression was used to analyze potential relationships between the independent variables and the dependent variable to determine if any combination of them correctly predicted a default repayment status. Prior to performing the logistic regression analyses, the data were reviewed for missing values and outliers. Descriptive statistics were obtained for the overall population, as well as for each of the three generational groups under investigation. Finally, the logistic regression was completed and compared to the null hypothesis.

Data Review and Transformation

Logistic regression requires a substantial amount of interpretation. However, this technique does not require the researcher to assume or verify a normal distribution within the variables, equal variances within each group, or that linear relationships exist between the variables (Courvoisier et al., 2010; Mertler & Vannatta, 2010). The flexibility of the analysis permits non-linear relationships among the variables, abnormal distributions, and unequal variances. Logistic regression is sensitive to outliers in the data and high correlations among variables (Anderson, Jin, & Grunkemeier, 2003; Mertler & Vannatta, 2010). Because of the sensitivity of this multivariate statistical technique, categorical variables were converted into numerical format and each variable was screened carefully for missing data, outliers, and linearity.

Missing Data and Recoding

Initial review of the data showed no vacant or incomplete data points within the data. When recoding the categorical variables of repayment status, gender, and enrollment status into a numeric format, there were two cases identified in which gender was reported as "N" for no response. Those two cases were removed from the data set and excluded from the analyses. An additional two cases were removed from the data set because the date of birth was outside of the scope of the study. Table 2 illustrates how the dependent and two independent categorical variables were recoded into numerical format.

Table 2

Recoding of Categorical Variables

Variable	Measurement	Coding
Loan Repayment Status	Categorical	0 = Non-default, 1 = Default
Gender	Categorical	0 = Male, 1 = Female
Enrollment Status	Categorical	0 = Graduated, 1 = Withdrawn, 2 = Other

Outliers and Linearity

As stated, logistic regression analysis is sensitive to outliers and a high correlation (multicollinearity) among the independent variables. As with outliers, multicollinearity within the variables can result in a higher weight being placed on two similar categories in the model when, in effect, they represent the same thing (Anderson, Jin, & Grunkemeier, 2003; Sprinthall, 2007). A preliminary multiple regression analysis assisted in the identification of outliers and multicollinearity among the independent variables. The output from the multiple regression analysis included the Mahalanobis distance test and the table of coefficients that lists tolerance values for each of the variables.

Outliers. The Mahalanobis distance test identifies outliers within the data. When running the multiple regression analysis, a new variable is created that provides each case's Mahalanobis distance value within the data set. The residuals statistics Table 3 provides the Mahalanobis distance minimum and maximum values.

Table 3

Table of Residuals Statistics to Identify Outliers

	Minimum	Maximum	M	SD	N
Predicted value	.04	.25	.14	.048	2290
Std. predicted value	−2.113	2.355	.000	1.000	2290
Standard error of predicted value	.010	.050	.017	.006	2290
Adjusted predicted value	.04	.26	.14	.048	2290
Residual	−.254	.939	.000	.343	2290
Std. residual	−.738	2.732	.000	.999	2290
Stud. residual	−.741	2.738	.000	1.000	2290
Deleted residual	−.256	.943	.000	.344	2290
Stud. deleted residual	−.741	2.741	.000	1.001	2290
Mahal. distance	.920	46.831	4.998	4.504	2290
Cook's distance	.000	.018	.000	.001	2290
Centered leverage value	.000	.020	.002	.002	2290

The critical value of chi-square χ^2 at $p < .001$ and degrees of freedom equal to 4 is 18.47. Consequently, cases with a Mahalanobis distance greater than 18.47 are considered multivariate outliers for the variables of gender, total debt, total lenders, enrollment, and year of birth. This process resulted in the exclusion of 36 cases in which their Mahalanobis distance value exceeded the chi-square value of 18.47 indicating that the cases were outliers within the established data set.

Linearity. The table of coefficients tolerance values indicates whether there is a high correlation between any of the variables. If multicollinearity were a factor, one of the variables would be excluded from the study to limit duplicative variables from affecting analysis results. Table 4 lists the tolerance values for each of the variables included in the study.

If the tolerance values are greater than .1 for each of the variables, no violation of multicollinearity between the particular variables is assumed (Menard, 2011; Tabachnick & Fidell, 2007). As each of the tolerance values provided were greater than .1, there was no violation of multicollinearity and no variables were excluded from analysis.

Table 4

Table of Coefficients[a]

Model		Collinearity Statistics	
		Tolerance	VIF
1	Gender3	.959	1.042
	Enroll_3	.973	1.028
	Lenders3	.667	1.500
	Year of birth	.908	1.101
	Total debt	.615	1.626

a. Dependent variable: Borrow_Stat

Descriptive Statistics

A series of descriptive statistics were conducted to obtain baseline data on the population as a whole, as well as for each of the three generations under review. The individual frequency information for each variable, average debt, and age of the entire population are provided in Appendix B. Table 5 provides generic descriptive information pertinent to the population under investigation.

The descriptive statistics of the population show that approximately 66% of the borrowers are female. Additionally, 80% of the population had a last reported enrollment status of withdrawn from the institution. The majority of borrowers in repayment also had two or fewer loan servicers. Of the borrowers in repayment, 57% are part of the millennial generation with another 36% within the Generation "X" group. Generation-specific tables providing the frequency for each variable are provided in appendices C-E.

Table 5

Population Frequency and Percentage

Enrollment	*f*	%	Generation	F	%	Servicer	F	%
Graduated	276	12.1	Millennial	1308	57.1	1	1641	71.7
Withdrawn	1832	80	Gen "X"	831	36.3	2	546	23.8
Other	182	7.9	Baby boomer	151	6.6	3	94	4.1
Repay Status			Gender			4	7	.3
Non-default	1970	86	Male	774	33.8	5	2	.1
Default	320	14	Female	1516	66.2			

Stratified Random Sampling

The population provided from the host institution included 2292 total cases. After outliers and missing data values were removed, there were 2252 remaining viable cases for sampling. Stratified random sampling was conducted based on gender and loan repayment status to keep the two sample groups representative of the overall population group.

Research suggests that to limit model bias, a minimum of 10 to 20 events per variable is recommended (Courvoisier et al., 2010; Peng, Lee, & Ingersoll, 2002). At 6 total variables and 20 events per variable, the researcher needed a minimum of 120 cases for each logistic regression analysis to limit model bias. By pulling 20 events per variable, the sample contained an effect size of 0.15 with a minimum 0.80 power size.

This is important to note because the sample size needs to be large enough to limit type one errors where the subsequent model is over-fitted to the sample group. In larger sample groups where the power size is too large, the probability of type two errors increases as the sample becomes unmanageable and contains too many variances, thereby, reducing the statistical significance of any resulting model (Peng, Lee, & Ingersoll, 2002).

From the population of individuals in loan repayment, Microsoft Excel was used to create two stratified random sample groups for each of the tests highlighted in research question one. The reason to pull two sample groups was to improve the predictive accuracy of the research question and subsequent model(s) (Anderson, Jin, & Grunkemeier, 2003). See Table 6 below for a breakdown of the stratification values used to develop the two sample groups.

Table 6

Random Sample Stratification Values

Repayment Status	Men	Women	Total
Non-default	33	70	103
Default	8	9	17
Total	41	79	120

Logistic Regression Outcomes

Direct logistic regression was performed twice to assess the impact of a number of factors on the likelihood that borrowers would default on their student loan debt. Both models contained five independent variables (i.e., gender, year of birth, last reported enrollment at the institution, total debt, and total number of loan servicers).

Sample Group One

The full model containing all predictors was not statistically significant when using the first sample group. The chi-square values, χ^2 (5, N = 120), = 10.140, $p < .001$, indicated that the first sample model was not able to distinguish between borrowers who would and would not default on their student loan debt. The model as a whole explained between 8.1% (Cox and Snell R^2) and 14.5% (Nagelkerke R^2) of the variance in repayment status, and correctly classified 87.5% of cases. This was a 1.7% improvement from the null hypothesis in which none of the variables was included in the model.

Table 7

Sample One: Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B) LL	95% C.I. for EXP(B) UL
Step 1[a]	lenders3	1.244	.557	4.982	1	.026	3.470	1.164	10.346
	gender3(1)	-.587	.565	1.080	1	.299	.556	.184	1.682
	enroll3			.740	2	.691			
	enroll3(1)	.978	1.137	.740	1	.390	2.659	.286	24.699
	enroll3(2)	-18.588	13223.188	.000	1	.999	.000	.000	.
	birthyear	-.006	.036	.030	1	.862	.994	.927	1.066
	totaldebt	.000	.000	.674	1	.412	1.000	1.000	1.000
	Constant	8.779	70.605	.015	1	.901	6496.565		

a. Variable(s) entered on step 1: lenders3, gender3, enroll3, birthyear, totaldebt

As shown in Table 7, only one of the independent variables made a unique statistically significant contribution to the model (number of lenders). The number of lenders was the strongest predictor of reporting a default repayment status, recording an odds ratio of 3.47. This indicated that borrowers with more than one loan servicer were more than three times more likely to default on their student loans than those borrowers with only one loan servicer were.

Sample Group Two

The full model containing all predictors was not statistically significant when using the second sample group. The chi-square values, χ^2 (5, N = 120), = 11.201, p < .001, indicated that the second sample model was not able to distinguish between borrowers who would and would not default on their student loan debt. The model as a whole explained between 8.9% (Cox and Snell R^2) and 16% (Nagelkerke R^2) of the variance in repayment status and correctly classified 85.8% of cases. This was no change from the null hypothesis in which none of the variables was included in the model.

As shown in Table 8, none of the independent variables was statistically significant to contribute to the model. Of note though, is that a withdrawn enrollment status appears to have the strongest predictor of reporting a default repayment status, recording an odds ratio of 4.03. This indicated that borrowers who withdraw from the institution were more than four times likely to default on their student loans than those borrowers with only one loan servicer. However, this category was not found to be statistically significant to predict loan repayment status, and caution should be used when interpreting the odds ratio for enrollment.

Table 8

Sample Two: Variables in the Equation

								95% C.I. for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	gender3(1)	-.693	.615	1.272	1	.259	.500	.150	1.668
	enroll3			.000	2	1.000			
	enroll3(1)	19.814	14867.030	.000	1	.999	4.028E8	.000	
	enroll3(2)	.280	18358.236	.000	1	1.000	1.323	.000	
	lenders3	-1.009	.699	2.084	1	.149	.365	.093	1.435
	birthyear	.017	.043	.150	1	.698	1.017	.934	1.107
	totaldebt	.000	.000	2.113	1	.146	1.000	1.000	1.000
	Constant	-53.678	14867.279	.000	1	.997	.000		

a. Variable(s) entered on step 1: gender3, enroll3, lenders3, birthyear, totaldebt.

Tables 9 and 10 represent side-by-side comparisons of models 1 and 2 for the test of coefficients and for the variance in loan repayment. The tables show that neither sample group was able to produce a statistically significant model of loan repayment status.

Table 9

Omnibus Tests of Model Coefficients

Group	Chi-square	Df	Sig.
Sample 1	10.140	6	.119
Sample 2	11.201	6	.082

Table 10

Models Summary for Variance in Loan Repayment

Sample	−2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	87.775[a]	.081	.145
2	86.714[a]	.089	.160

a. Estimation terminated at iteration number 20 because maximum iterations have been reached. The final solution cannot be found.

Based on the resulting outcomes where $p > .005$ for both sample tests, a statistically significant and reliable predictive model was not created. As a result, this research did not reject the null hypothesis.

CHAPTER FIVE: CONCLUSIONS

Tenacity is the underlying thread for all research. Without a healthy dose of inquiry and persistent determination, answers to large and often complex questions would likely go unanswered. Student loan repayment is a complex area that intertwines a host of borrower, institutional, economic, and socially-specific characteristics.

As stated before, relevant research in the area of student loan default is outdated with a number of studies more than a decade old. Findings from those researchers are appropriate for a different social and economic context than what is relevant today. The findings from this study incorporate a new generation—the millennials, who were previously missing from prior studies. Additionally, findings from this study utilize 2013 data giving the institution access to immediate information regarding their loan cohort, borrower trends, and default characteristics.

Discussion of Results

This study attempted to develop a predictive model of student loan default from the variables of gender, age, number of loan servicers, last reported enrollment status, and total undergraduate loan debt. The independent variables selected were shown to be strong predictors of student loan default in previous studies as mentioned earlier. The researcher employed stratified random sampling based on gender and loan repayment status to develop two sample groups. The study required two separate logistic regression analyses to be conducted on two independent stratified random samples of the population. The research design was to develop a first initial model with the first logistic regression analysis, and then attempt to validate it with a second logistic regression analysis.

The results of the first regression analysis showed that none of the independent variables reliably predicted a borrower's loan repayment status. The second regression analysis confirmed the results of the first analysis in that no reliable or accurate predictive model could be developed with the variables under investigation. Of particular note was that the findings of the first logistic regression analysis showed the odds of default increase by a factor of 3.47 for borrowers who have more than one loan servicer. No other remarkable statistically significant findings were found from the results of the logistic regression analyses.

Conclusion

In conclusion, a predictive model of student loan default could not be developed from this study. The findings from this study are unique compared to relative other available research within the field. The reasons for this could be the fact that relevant literature within the field is approximately two decades old, with more recent research looking at four-year institutions. This study focused and obtained data from a large two-year community college on the East Coast. The lack of a predictive model could be unique to the specific institution and not necessarily relative to the population as a whole.

Another item for discussion is that this study examined and included millennial generation borrowers, which has not been a highly studied group. Available research on millennial generation borrowers and their repayment trends is limited (Cunningham & Kienzel, 2011). Independent variables were shown to be strong predictors in previous research for "X" and baby boomer generations (Christman, 2000; Flint, 1997; Gross, et al., 2009; Kesterman, 2005; Steiner & Teszler, 2003). As the millennial generation is a relatively unstudied population, variables proven in previous research may not be relative to the majority generation enrolled in college today. This study showed that variables likely to predict default for previous generations showed to be ineffective at predicting default from a mixed sample where millennial borrowers constituted 57% of the population.

Theoretical Inferences

Without the development of a predictive model, inferring a theory about why borrowers default on their loan debt could prove to be a difficult challenge. While specific theories cannot be directly drawn upon to explain a predictive model, theoretical foundations were inferred from the data set.

Economic

The first logistic regression analysis showed that students with two or more loan servicers were 3.47 times more likely to default on their student loan debt. Both the human capital theory and debt burden theory could be applied toward this correlation.

The human capital theory posits that students will invest in themselves when there is the potential for long-term financial gains (Becker, 1994). Students with two or more servicers could be indicative of the borrower's willingness to invest or finance their education with the expectation that their annual and lifetime earnings will increase.

When the number of loan servicers increases so too does the borrower's total debt. Debt burden theory hypothesizes that borrowers are able to pay only those bills in which the debt does not become unmanageable (Cunningham & Kienzel, 2011; Kesterman, 2005). The American Council on Education (2004), places that manageable debt threshold at 8% of monthly income. Student loan debts that exceed 8% of the monthly income are subsequently viewed as unmanageable under the tenants of this theory. Borrowers with two or more servicers would indicate that the borrower has at minimum two or more loans in repayment. If that total debt exceeds the 8% threshold, the amount could be unmanageable to the borrower, resulting in an inability to pay the monthly payments.

Sociological

The approach-avoidance theory can also be used to help explain why borrowers with more than one loan servicer may be more likely to default than their counterparts with one servicer. In this situation, the approach-avoidance theory would posit that borrowers become overwhelmed when faced with two or more servicers contacting them for payment. Instead of demonstrating a positive response to work with the lenders, the borrowers may demonstrate a negative response and avoid repayment with one or more servicers.

The number of servicers a borrower has also creates an atmosphere in which confusion can develop as to which servicer holds what loans, from what institutions, and for how much. This confusion can compound if the loan servicers are in the middle of selling or transferring portions of their loan portfolios to other companies. As loans are in transition from one corporate portfolio to another, borrowers may become frustrated or confused about who is contacting them for payment and subsequently stop payments altogether.

A final note from the demographic statistics of the population was that 80% of the borrowers had a last reported enrollment status of withdrawn. Tinto (1987) theorized that the reasons for student departure were results from the degree-of-fit of the experiences the student had while at the institution. Borrower's who were not able to connect with the institution or possessed negative beliefs about the institution or education were likely to withdraw from their program of study prior to completion (Cabrera, Stampen, & Hansen, 1988; Coleman, 2010; Jenkins & Lynch, 2007). Community colleges often struggle with graduation rates because students will transfer to other universities to complete their degrees, transition back into the workforce, or withdraw based on family, social, or economic situations (Chen, 2015).

Implications for Practice

Findings from this research left more questions than provided answers. The relevant variables from prior research were not statistically significant with this data set. One model found students to be 3.47 times more likely to default on their student loan debt as the number of loan servicers increased. No viable predictive model of student loan default was created at the institutional level. As a result, no model could be extrapolated for possible state or national consideration.

This research has provided a general idea of the borrower population that attends a large community college on the East Coast. Demographics show that two-thirds of the borrower populations are women and that 80% of the population had a last reported enrollment status of withdrawn.

The research also provides a framework for future studies to investigate student loan repayment patterns at other institutions. This research adds an important component into the student loan repayment equation by including millennial generation borrowers in the analyses. Previous research is outdated and has left the millennial generation out of consideration. The reason for this is that the millennial generation had yet to enter higher education and start loan repayment. By including millennial borrowers in the study, this research provides a varied and comprehensive look at today's diverse borrower demographics. The findings can assist concerned stakeholders in the development of timely financial literacy programs and enhanced default management outreach activities.

Recommendations for Future Research

Two questions loom over this study. The first being whether the results are specific to the institution. The second is whether the results are the outcome of a complex and diverse millennial generation that is still relatively unexplored concerning loan repayment behaviors. The variables that showed to be positive indicators of default for the baby boomer and generation "X" age groups are not viable for this new generation of borrowers. The resulting question then is: What characteristics help predict student loan default among a diverse multi-generation population? What characteristics help predict student loan default specific to the millennial generation?

Millennial Generation

Additional comprehensive studies that specifically look at repayment patterns and trends of the millennial generation in various higher education settings would help highlight the traits relative to the millennial borrower. The findings from this line of research could also help identify any differences in nuances between the generations.

Multi-Generational

As society changes and develops over time, so do borrower demographics, technology, and the portrait of higher education students. Understanding the diverse needs of multiple generations will assist stakeholders in identifying training requirements of staff, educational needs of future and returning borrowers, and repayment assistance information to those entering into repayment.

Institutional Type

As previously mentioned, an unknown for this study is whether the findings are specific to the institution or commonly seen at the statewide or national level. Do these findings change with borrowers who attend a four-year public university? What predictive model of student loan repayment might exist for two-year private institutions? Expanding the research by examining the various institutional types and sizes can assist in facilitating a better understanding of whether the findings of this study are institution-specific or if this is the new norm for most of the national population.

Institution-Specific Variables

A further area for future research would be to expand the institution-specific variables. In addition to last reported enrollment status, expanding institutional variables to include fields such as major, program of study, time to completion of program, college grade point average, among other characteristics may paint a better picture of repayment trends. If variables are found at the institutional level that predicts student loan repayment status, then schools could monitor their borrower population better. This institution-specific information could assist stakeholders in the development of financial literacy counseling and default management strategies to educate and reduce the likelihood of borrower default.

Borrower-Specific Variables

A final area of further research would be to expand borrower-specific variables to explore what additional variables may be predictive of loan repayment. These variables can include, but are not exclusive to, borrower ethnicity, household size, familial, educational background, financial ability, and household income. Understanding the specific nuances of the individual borrower and identifying potential barriers to repayment can further assist stakeholders in developing default management and financial literacy programs.

REFERENCES

American Council on Education (2004, September). *Debt burden: Repaying student debt.* Retrieved October 15, 2012, from http://www.acenet.edu/programs/policy

Anderson, R. P., Jin, R., & Grunkemeier, G. L., (2003). Understanding logistic regression analysis in clinical reports: An introduction. *The Annals of Thoracic Surgery*, *75*, 753-757. Retrieved from http://ats.ctsnetjournals.org

Barone, S., Steiner, M., & Teszler, N. (2005, May). *Multivariate analysis of student loan defaulters at Texas A&M University – Kingsville*. Retrieved from Texas Guarantee Student Loan Corporation: http://www.tgslc.org/

Baum, S., & O'Malley, M. (2003). College on credit: How borrowers perceive their education debt. Results of the 2002 National Loan Survey. *Journal of Student Financial Aid*, *33*(3), 7-19. Retrieved from http://www.eric.ed.gov/

Baum, S., & Schwartz, S. (2005, November). *How much debt is too much? Defining benchmarks for manageable student debt.* Retrieved October 15, 2012, from ProQuest database.

Becker, G. S. (1994). *Human capital: A theoretical and empirical analysis with special reference to education* (3rd ed.). Retrieved from http://www.nber.org/books/beck94-1

Becker, G. S., & Mulligan, C. B. (1997). The endogenous determination of time preference. *Quarterly Journal of Economics*, (3), 729-758. doi:10.1162/003355397555334.

Bottone, G., & Sena, V. (2011). Human capital: Theoretical and empirical insights. *American Journal of Economics and Sociology*, *70*(2), 401-423. Retrieved from ABI/INFORM Complete database.

Boyd, L. A. (1997). Discrimination in mortgage lending: The impact on minority defaults in the Stafford Loan program. *The Quarterly Review of Economics and Finance*, *37*(1), 23-37. Retrieved from http://papers.ssrn.com/

Budget Control Act, 2 U.S.C. § 900. (2011). United States 112th Congress.

Cabrera, A. F., Stampen, J. O., & Hansen, W. L. (1988, November 3-6). *Exploring the effects of ability-to-pay on persistence in college* (Report No. ED303096). Retrieved from Education Resources Information Center database.

California Postsecondary Education Commission. (2006). *Developing a statewide higher education affordability policy* (Commission Report 06-10). Sacramento, CA: CPEC.

Catalyst. (2012). *Generations in the workplace in the United States & Canada*. Retrieved January 8, 2013, from http://www.catalyst.org/

Chen, G. (2015, March 3). *The catch-22 of community college graduation rates*. Retrieved March 30, 2015 from http://www.communitycollegereview.com

Chopra, R. (2012, March 21). *Too big to fail: Student debt hits a trillion* [White paper]. Retrieved from Consumer Financial Protection Bureau: http://www.consumerfinance.gov/blog/too-big-to-fail-student-debt-hits-a-trillion/

Choy, S. P., & Li, X. (2006). *Dealing with debt: 1992-93 bachelor's degree recipients 10 years later* (NCES 2006-156). Washington, DC: Us Department of Education, National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2006/2006156.pdf

Christman, D. E. (2000). Multiple realities: Characteristics of loan defaulters at a two-year public institution. *Community College Review*, *27*(4), 16-32. Retrieved from ProQuest database.

Clarke, R. J. (2005). *HDR seminar series: Research methodologies* [PowerPoint slides]. Retrieved from HDR Seminar Series—Faculty of Commerce at: http://www.uow.edu.au/

"Cohort Default Rates," Title 34 Code of Federal Regulations Pt. 668, Subpart N. 2015 ed.

Coleman, E. (2010). *Money matters: A study on the relationship between financial aid programs and policies on student persistence in higher education*. Retrieved from ProQuest. (UMI 3422657).

Conference Board of Canada. (2009). *Winning the generation wars: Making the most of generational differences and similarities in the workplace.* Retrieved from https://www.aqesss.qc.ca/docs/pdf/i-media/20091126/ConferenceBoard_Compete_Generation_Wars.pdf

Council for Higher Education Accreditation. (2010). *U.S. Department of Education publishes proposed regulations addressing program integrity and student aid programs.* Retrieved from http://www.chea.org/

Courvoisier, D., Combescure, C., Agoritsas, T., Gayet-Ageron, A., & Perneger, T. (2010). Performance of logistic regression modeling: Beyond the number of events per variable, the role of data structure. *Journal of Clinical Epidemiology,* 64(2011), 993-1000. doi: 10.1016/j.jclinepi.2010.11.012

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cunningham, A. F., & Kienzel, G. S. (2011). *Delinquency: The untold story of student loan borrowing*. Retrieved from www.ihep.org

Department of Education: Federal student aid. (2010). *Update on student loan programs*. Retrieved October 19, 2011, from http://studentaid.ed.gov/

Department of Education: Federal student aid. (2011a). *National student loan default rates* [Fact sheet]. Retrieved from http://www2.ed.gov/offices/OSFAP/ defaultmanagement/defaultrates.html

Department of Education: Information for financial aid professionals. (2011b). *Systems & processing links—application*. Retrieved July 23, 2011, from http://ifap.ed.gov

Department of Labor. (2011). *Unemployment rates*. Retrieved May 15, 2011, from www.dol.gov

Eaton, S. B., & Bean, J. P. (1995). An approach/avoidance behavioral model of college student attrition. *Research in Higher Education*, *36*(6), 617-645. Retrieved from ProQuest Central database

Federal student aid. (n.d.). *Default Management*. Retrieved May 4, 2015 from http://www.ifap.ed.gov/DefaultManagement/DefaultManagement.html

Federal student aid. (2010a). *FY2007 cohort default rates*. Retrieved July 21, 2011, from http://federalstudentaid.ed.gov/datacenter/cohort.html

Federal student aid. (2010b). *National student loan default rates*. Retrieved July 21, 2011, from http://www2.ed.gov/offices/OSFAP/defaultmanagement/defaultrates. html

Federal student aid (2014). *National student loan default rates*. Retrieved August 1, 2014, from http://www2.ed.gov/offices/OSFAP/defaultmanagement/defaultrates.html

Flint, T. A. (1997). Predicting student loan defaults. *The Journal of Higher Education*, *68*(3), 322-354. Retrieved from ProQuest database.

Fraas, C. J. (1989, July 31). *The U.S. Department of Education's student loan default reduction initiative: Background and analysis* [White paper]. Retrieved from Education Research Information Center: http://www.eric.ed.gov

Gelston, S. (2008). *Gen Y, gen X, and the baby boomers: Workplace generation wars*. Retrieved January 13, 2013, from http://www.cio.com.au/

General Accounting Office. (1991). *Student loans: Characteristics of defaulted borrowers in the Stafford student loan program*. Retrieved from http://www.eric.ed.gov

Gladieux, L., & Perna, L. (2005, May). *Borrowers who drop out* (National Center Report #05-2). Retrieved from The National Center for Public Policy and Higher Education: http://www.highereducation.org/reports/borrowing/index.shtml

Goodwin, L. (2011). *Fact check: The exploding student loan debt chart*. Retrieved December 14, 2011, from http://news.yahoo.com/blogs/lookout/fact-check-exploding-student-loan-debt-chart-145504148.html

Gross, J. P., Cekic, O., Hossler, D., & Hillman, N. (2009). *What matters in student loan default: A review of the research literature*. Retrieved May 27, 2011, from www.nasfaa.org

Harrast, S. A. (2004). Undergraduate borrowing: A study of debtor students and their ability to retire undergraduate loans. *Journal of Student Financial Aid*, *34*(1), 21-37. Retrieved from http://www.eric.ed.gov/

Herr, E., & Burt, L. (2005). Predicting student loan default for the University of Texas at Austin. *Journal of Student Financial Aid*, *35*(2), 27-49. Retrieved from http://www.nasfaa.org/research/Journal/Journal_of_Student_Financial_Aid.aspx

Higher Education Reconciliation Act, 34 C.F.R. § Part 668, Subparts M and N *et seq.* (1965). United States.

Jenkins, M. P., & Lynch, J. M. (2007). *Report on student debt and alumni giving*. Retrieved from American Student Assistance: http://www.asa.org

Kantrowitz, M. (2009). *Characteristics of borrowers with excessive debt.* Retrieved from http://www.finaid.org

Kesterman, F. (2005). *Student loan borrowing in America: Metrics, demographics, and loan default aversion strategies*. Retrieved from ProQuest Dissertations and Theses. (UMI Number: 3156011).

Kypuros, C. A. (2009). *Factors affecting student loan default rates: Nevada System of Higher Education*. Retrieved from ProQuest Dissertations and Theses. (3383982).

Lazarus, R., Averill, J., & Opton, E. (1974). The psychology of coping: Issues of research and assessment. In G.V. Coelho, D. A. Hamburg and J. E. Adams (Eds.), *Coping and Adaptation*. New York, NY: Basic Books, 1974, 249-315.

Lederman, D. (2011). *Trouble ahead on student loan defaults*. Retrieved October 19, 2011, from http://www.insidehighered.com/news/2011/05/23/student_loan_default_rates_rise_sharply_especially_for_for_profit_colleges

Lee, J. (2009). *Racial/ethnic disparities in household debt repayment*. Retrieved from Proquest LLC database. (UMI Number: 3375793).

Leimkuhler, F. F. (1972). Library operational research: A process of discovery and justification. In Swanson, D.R., & Bookstien, A. (Eds.), *Operations research*. (pp. 84-96). University of Chicago.

Lochner, L., & Monge-Naranjo, A. (2003). *Education and default incentives with government student loan programs*. Retrieved July 8, 2013, from http://citeseerx.ist.psu.edu/

Lochner, L., & Monge-Naranjo, A. (2004). *Education, credit constraints and default incentives*. Retrieved July 8, 2013, from http://faculty.wcas.northwestern.edu/

Loonin, D., & McLaughlin, J. (2012, July). *The student loan default trap: Why borrowers default and what can be done*. Retrieved from National Consumer Law Center Web site: www.nclc.org

Lund Research. (2012). *Stratified random sampling*. Retrieved April 3, 2013, from http://dissertation.laerd.com/

Menard, S. (2011). Standards for standardized logistic regression coefficients. *Social Forces*, 89, 1409-1428.

Mertler, C. A., & Vannatta, R. A. (2010). *Advanced and multivariate statistical methods* (4th ed.). Glendale, CA: Pyrczak Publishing.

Mishory, J., & O'Sullivan, R. (2012). *Denied? The impact of student debt on the ability to buy a house*. Retrieved October 26, 2012, from http://younginvincibles.org/

Missouri Department of Higher Education. (n.d.). *Financial literacy*. Retrieved July 23, 2011, from http://dhe.mo.gov/

Mullins, E. (2008). *Debt and decision-making: The impact of student loan debt on undergraduates' choice of major, vocation, and post-college plans*. Retrieved from ProQuest Dissertations and Theses. (UMI Number: 1475722).

National Association of Student Financial Aid Administrators. (2011). *Deficit committee dismantles triggering automatic cuts in FY 2013*. Retrieved November 22, 2011, from http://www.nasfaa.org

National Student Loan Database System. (2012). *FAQs*. Retrieved November 26, 2012, from https://www.nslds.ed.gov

National Summit of Councils. (2001). *National summit of councils: Focusing on financial aid management*. Retrieved July 19, 2011, from www.tgslc.org

New America Foundation. (2013). *Federal education budget project: Federal Pell grant program*. Retrieved July 5, 2013, from http://febp.newamerica.net/

Office of Scholarships and Financial Aid: Loyola University New Orleans. (2011). *Default rate statistics*. Retrieved October 19, 2011, from http://www.loyno.edu/financialaid/default-rate-statistics

Pascarella, E. E. (1985). Students' affective development within the college environment. *Journal of Higher Education*, *56*(6), 640-663. Retrieved from http://www.jstor.org/

Peng, C. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. *The Journal of Educational Research*, *96*(1). doi:00220671.

Pierson, J., Walsh, M., & Wright, R. (2011). *Delinquency management and default prevention* [PowerPoint slides]. Retrieved from http://www2.ed.gov/offices/OSFAP/training/downloads.html

Podgursky, M., Ehlert, M., Monroe, R., Watson, D., & Wittstruck, J. (2002). Student loan defaults and enrollment persistence. *Journal of Student Financial Aid*, *32*(3), 27-42. Retrieved from http://www.nasfaa.org/

"Program Integrity Issues" 75 Federal Register 209 (2010, October), pp. 66832-6684

Salas-Amaro, A. (2008). *Default prevention and management practices at Florida community colleges*. Retrieved from ProQuest Dissertations and Theses. (UMI 3302626).

Schwartz, S., & Finnie, R. (2002). Student loans in Canada: An analysis of borrowing and repayment. *Economics of Education Review, 21*(5), 497-512.

Shafique, F., & Mahmood, K. (2010). *Model development as a research tool: An example of PAK-NISEA*. Retrieved November 21, 2013, from http://www.webpages.uidaho.edu/~mbolin/shafique-mahmood.htm

Smith, E. J. (2002). *Policies and procedures of California community colleges related to student loan default rates.* Retrieved from ProQuest Dissertations and Theses. (UMI Number: 3086880).

Sprinthall, R. C. (2007). *Basic statistical analysis,* (8th ed.). Boston, MA: Pearson.

Steiner, M., & Teszler, N. (2003, January). *The characteristics associated with student loan default at Texas A&M University*. Retrieved from Texas Guarantee Student Loan Corporation: http://www.tgslc.org/pdf/tamu_default_study.pdf

Stockburger, D. W. (1996). *Introductory statistics: Concepts, models, and applications.* Retrieved on November 26, 2013, from http://www.psychstat.missouristate.edu/introbook/sbk04m.htm

Student Aid and Fiscal Responsibilities Act, 42 U.S.C. § 1305. (2010). 111th Congress.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston, MA: Pearson/Allyn & Bacon.

Theilfodlt, D., & Scheef, D. (2004). *Generation X and the millennials: What you need to know about mentoring the new generations*. Retrieved January 13, 2013, from http://apps.americanbar.org/

Tinto, V. (1987, November 20). *The principles of effective retention* (ED301267). Retrieved from Education Resources Information Center (ERIC): ERIC database

"Two Year Cohort Default Rates," Title 34 Code of Federal Regulations Pt. 668, Subpart M. 2015 ed.

U.S. Department of Education. (2013). *Default prevention and management*. Retrieved March 28, 2013, from http://ifap.ed.gov

U.S. Equal Employment Opportunity Commission. (n.d.). *Questions and answers - Implementation of revised race and ethnic categories* [Fact sheet]. Retrieved from http://www1.eeoc.gov//employers/eeo1/qanda-implementation.cfm?renderforprint=1

Volkwein, J. F., & Szelest, B. P. (1995). Individual and campus characteristics associated with student loan default. *Research in Higher Education*, *36*(1), 41-72.

Volkwein, J. F., Cabrera, A. F., Szelest, B. P., & Napierski, M. R. (1998). Factors associated with student loan default among different racial and ethnic groups. *Journal of Higher Education*, *69*(2), 206-237. Retrieved from Education Resources Information Center database.

Woo, J. H. (2002). *Clearing accounts: The causes of student loan default* [White paper]. Retrieved from EdFund: http://www.edfund.org/

APPENDICES

APPENDIX A

2011 Three-Year CDR Report Variables Defined

APPENDIX A

2011 Three-Year CDR Report Variables Defined

Table A1

2011 Three-Year CDR Report Variables Defined

Variable Provided	Variable Definition
PIDM	Unique random identifier
D.O.B.	Date of birth
Orig sch	Originating school
Ind	Indicator
Begin date	Loan begin date
End date	Loan end date
Academic level	Academic grade level
Lender/holder current	Current lender/owner of the loan
Loan type	Federal direct loan or FFELP subsidized or unsubsidized
Stat	Repayment status
Date	Date of repayment status
Claim rsn/code	Reason code to explain specific loan repayment statuses
Default/negam date	Date of default
Repay date	Repayment date
Amount	Amount of loan
Guarantor/servicer	Guarantor/loan servicer responsible for collecting on loan debt
Guaranty loan date	Date loan was guaranteed
Enrollment code	Last reported enrollment status at the institution
Stat date usage1	For use in the CDR calculation indicating numerator or denominator
Gender	Gender of the borrower

APPENDIX B

Population Demographics and Relevant Statistics

APPENDIX B

Population Demographics and Relevant Statistics

Table B1

Population Enrollment Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Graduated	276	12.1	12.1	12.1
Withdrawn	1832	80.0	80.0	92.1
Other	182	7.9	7.9	100.0
Total	2290	100.0	100.0	

Table B2

Population Gender Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	774	33.8	833.8	33.8
Female	1516	66.2	66.2	100.0
Total	2290	100.0	100.0	

Table B3

Population Borrower Status Frequency and Percentages

	Frequency	Percent	Valid Percent	Cumulative Percent
Denominator[a]	1970	86.0	86.0	86.0
Both[b]	320	14.0	14.0	100.0
Total	2290	100.0	100.0	

a. Unduplicated borrowers of the 2011 cohort who are not in default.
b. Unduplicated borrowers of the 2011 cohort who are in default.

Table B4

Number of Loan Servicers per Borrower Frequency and Percentages

Servicer	Frequency	Percent	Valid Percent	Cumulative Percent
1	1641	71.7	71.7	71.7
2	546	23.8	23.8	95.5
3	94	4.1	4.1	99.6
4	7	.3	.3	99.9
5	2	.1	.1	100.0
Total	2290	100.0	100.0	

Table B5

Population Generation Membership Frequency and Percentages

Generation	Frequency	Percent	Valid Percent	Cumulative Percent
1982-2000	1308	57.1	57.1	57.1
1966-1981	831	36.3	36.3	93.4
1946-1965	151	6.6	6.6	100.0
Total	2290	100.0	100.0	

Table B6

Population Mean, Range, and Standard Deviation for Year of Birth

	N	Minimum	Maximum	Mean	Std. Deviation
Year of birth	2290	1947	1992	1981.23	8.532
Valid N (listwise)	2290				

Table B7

Population Mean, Range, and Standard Deviation for Total Debt

	N	Minimum	Maximum	Mean	Std. Deviation
Total debt	2290	149	49663	8136.76	6711.256
Valid N (listwise)	2290				

Table B8

Population Debt Frequency and Percentages in $5,000 Increments

Total Debt	Frequency	Percent	Valid Percent	Cumulative Percent
1-5000	921	40.2	40.2	40.2
5001-10000	814	35.5	35.5	75.8
10001-15000	251	11.0	11.0	86.7
15001-20000	155	6.8	6.8	93.5
20001-25000	84	3.7	3.7	97.2
25001-30000	33	1.4	1.4	98.6
30001-35000	13	.6	.6	99.2
35001-40000	6	.3	.3	99.4
40001-45000	7	.3	.3	99.7
45001-50000	6	.3	.3	100.0
Total	2290	100.0	100.0	

APPENDIX C

Millennial Generation Demographics and Relevant Statistics

APPENDIX C

Millennial Generation Demographics and Relevant Statistics

Table C1

Millennial Generation Independent Variable Validation

	Total Lenders	Enrollment	Gender	Borrow_Stat	Debt
Valid	1308	1308	1308	1308	1308
Missing	0	0	0	0	0

Table C2

Number of Loan Servicers for Each Millennial Borrower

	Frequency	Percent	Valid Percent	Cumulative Percent
1	979	74.8	74.8	74.8
2	283	21.6	21.6	96.5
3	41	3.1	3.1	99.6
4	4	.3	.3	99.9
5	1	.1	.1	100.0
Total	1308	100.0	100.0	

Table C3

Millennial Enrollment Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Graduated	126	9.6	9.6	9.6
Withdrawn	1078	82.4	82.4	92.0
Other	104	8.0	8.0	100.0
Total	1308	100.0	100.0	

Table C4

Millennial Gender Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	531	40.6	40.6	40.6
Female	777	59.4	59.4	100.0
Total	1308	100.0	100.0	

Table C5

Millennial Mean, Range, and Standard Deviation for Year of Birth and Total Debt

	N	Minimum	Maximum	Mean	Std. Deviation
Year of Birth	1308	1982	1992	1987.18	2.799
Total Debt	1308	149	38037	6450.73	4814.516
Valid N (listwise)	1308				

Table C6

Millennial Debt Frequency and Percentages in $5,000 Increments

	Frequency	Percent	Valid Percent	Cumulative Percent
1-5000	633	48.4	48.4	48.4
5001-10000	470	35.9	35.9	84.3
10001-15000	119	9.1	9.1	93.4
15001-20000	58	4.4	4.4	97.9
20001-25000	18	1.4	1.4	99.2
25001-30000	6	.5	.5	99.7
30001-35000	2	.2	.2	99.8
35001-40000	2	.2	.2	100.0
Total	1308	100.0	100.0	

Table C7

Millennial Borrower Status Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Denominator[a]	1108	84.7	84.7	84.7
Both[b]	200	15.3	15.3	100.0
Total	1308	100.0	100.0	

a. Unduplicated borrowers of the 2011 cohort who are not in default.
b. Unduplicated borrowers of the 2011 cohort who are in default.

APPENDIX D

Generation X Demographics and Relevant Statistics

APPENDIX D

Generation X Demographics and Relevant Statistics

Table D1

Generation X Independent Variable Validation

N	Total Lenders	Enroll_3	Gender3	Borrow_Stat	Debt3
Valid	831	831	831	831	831
Missing	0	0	0	0	0

Table D2

Number of Loan Servicers Per Generation X Borrower

	Frequency	Percent	Valid Percent	Cumulative Percent
1	565	68.0	68.0	68.0
2	218	26.2	26.2	94.2
3	45	5.4	5.4	99.6
4	2	.2	.2	99.9
5	1	.1	.1	100.0
Total	831	100.0	100.0	

Table D3

Generation X Enrollment Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Graduated	129	15.5	15.5	15.5
Withdrawn	638	76.8	76.8	92.3
Other	64	7.7	7.7	100.0
Total	831	100.0	100.0	

Table D4

Generation X Gender Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	208	25.0	25.0	25.0
Female	623	75.0	75.0	100.0
Total	831	100.0	100.0	

Table D5

Generation X Mean, Range and Standard Deviation for Year of Birth and Total Debt

	N	Minimum	Maximum	Mean	Std. Deviation
Year of birth	831	1966	1981	1975.71	4.233
Total debt	831	250	49663	10497.91	8203.192
Valid N (listwise)	831				

Table D6

Generation X Borrower Status Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Denominator[a]	729	87.7	87.7	87.7
Both[b]	102	12.3	12.3	100.0
Total	831	100.0	100.0	

a. Unduplicated borrowers of the 2011 cohort who are not in default.
b. Unduplicated borrowers of the 2011 cohort who are in default.

Table D7

Generation X Debt Frequency and Percentages in $5,000 Increments

	Frequency	Percent	Valid Percent	Cumulative Percent
1-5000	246	29.6	29.6	29.6
5001-10000	285	34.3	34.3	63.9
10001-15000	110	13.2	13.2	77.1
15001-20000	82	9.9	9.9	87.0
20001-25000	60	7.2	7.2	94.2
25001-30000	25	3.0	3.0	97.2
30001-35000	8	1.0	1.0	98.2
35001-40000	3	.4	.4	98.6
40001-45000	6	.7	.7	99.3
45001-50000	6	.7	.7	100.0
Total	831	100.0	100.0	

APPENDIX E

Baby Boomer Demographics and Relevant Statistics

APPENDIX E

Baby Boomer Demographics and Relevant Statistics

Table E1

Baby Boomer Independent Variable Validation

N	Total Lenders	Enroll	Gender	Borrow_Stat	Debt3
Valid	151	151	151	151	151
Missing	0	0	0	0	0

Table E2

Number of Loan Servicers per Baby Boomer Borrower

	Frequency	Percent	Valid Percent	Cumulative Percent
1	97	64.2	64.2	64.2
2	45	29.8	29.8	94.0
3	8	5.3	5.3	99.3
4	1	.7	.7	100.0
Total	151	100.0	100.0	

Table E3

Baby Boomer Enrollment Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Graduated	21	13.9	13.9	13.9
Withdrawn	116	76.8	76.8	90.7
Other	14	9.3	9.3	100.0
Total	151	100.0	100.0	

Table E4

Baby Boomer Gender Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Male	35	23.2	23.2	23.2
Female	116	76.8	76.8	100.0
Total	151	100.0	100.0	

Table E5

Baby Boomer Mean, Range, and Standard Deviation for Year of Birth and Total Debt

	N	Minimum	Maximum	Mean	Std. Deviation
Year of birth	151	1947	1965	1960.13	4.138
Total debt	151	500	43682	9747.37	7397.122
Valid N (listwise)	151				

Table E6

Baby Boomer Borrower Status Frequency and Percentage

	Frequency	Percent	Valid Percent	Cumulative Percent
Denominator[a]	133	88.1	88.1	88.1
Both[b]	18	11.9	11.9	100.0
Total	151	100.0	100.0	

a. Unduplicated borrowers of the 2011 cohort who are not in default.
b. Unduplicated borrowers of the 2011 cohort who are in default.

Table E7

Baby Boomer Debt Frequency and Percentages in $5,000 Increments

	Frequency	Percent	Valid Percent	Cumulative Percent
1-5000	42	27.8	27.8	27.8
5001-10000	59	39.1	39.1	66.9
10001-15000	22	14.6	14.6	81.5
15001-20000	15	9.9	9.9	91.4
20001-25000	6	4.0	4.0	95.4
25001-30000	2	1.3	1.3	96.7
30001-35000	3	2.0	2.0	98.7
35001-40000	1	.7	.7	99.3
40001-45000	1	.7	.7	100.0
Total	151	100.0	100.0	

APPENDIX F

Sample One Outputs and Data Tables

APPENDIX F

Sample One Outputs and Data Tables

Table F1

Case Processing Summary for Sample One

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	120	100.0
	Missing Cases	0	.0
	Total	120	100.0
Unselected Cases		0	.0
	Total	120	100.0

a. If weight is in effect, see classification table for the total number of cases.

Table F2

Dependent Variable Encoding

	Original Value	Internal Value
Dimension	D	0
	B	1

Table F3

Categorical Variables Coding

Variable	Code	Frequency	Parameter Coding (1)	Parameter Coding (2)
enroll 3	G	15	.000	.000
	W	96	1.000	.000
	Other	9	.000	1.000
gender 3	M	41	.000	
	F	79	1.000	

Block 0: Beginning Block

Table F4

Classification Table[a,b]

Observed			Predicted		
			Borrow Stat		Percentage Correct
			D	B	
Step 0	Borrow stat	D	103	0	100.0
		B	17	0	.0
	Overall percentage				85.8

a. The constant is included in the model.
b. The cut value is .500

Table F5

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.802	.262	47.357	1	.000	.165

Table F6

Variables not in the Equation[a]

			Score	df	Sig.
Step 0	Variables	lenders3	3.256	1	.071
		gender3(1)	1.464	1	.226
		enroll3	2.673	2	.263
		enroll3(1)	2.467	1	.116
		enroll3(2)	1.606	1	.205
		Birthyear	.136	1	.712
		totaldebt	.281	1	.596

a. Residual Chi-Squares are not computed because of redundancies.

Block 1: Method = Enter

Table F7

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	10.140	6	.119
	Block	10.140	6	.119
	Model	10.140	6	.119

Table F8

Sample One Model Summary

Step	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2
1	87.775a	.081	.145

a. Estimation terminated at iteration number 20 because maximum iterations have been reached. The final solution cannot be found.

Table F9

Sample One Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.537	8	.480

Table F10

Sample One Contingency Table for Hosmer and Lemeshow Test

		Borrow Stat = D		Borrow Stat = B		
		Observed	Expected	Observed	Expected	Total
Step 1	1	12	11.910	0	.090	12
	2	11	11.260	1	.740	12
	3	9	10.884	3	1.116	12
	4	11	10.777	1	1.223	12
	5	11	10.673	1	1.327	12
	6	11	10.592	1	1.408	12
	7	10	10.156	2	1.844	12
	8	9	9.844	3	2.156	12
	9	12	9.533	0	2.467	12
	10	7	7.370	5	4.630	12

Table F11

Sample One Model Classification Table[a]

Observed			Predicted		
			Borrow Stat		Percentage Correct
			D	B	
Step 1	Borrow stat	D	103	0	100.0
		B	15	2	11.8
	Overall percentage				87.5

a. The cut value is .500

Table F12

Sample One Variables in the Equation

Step 1a	B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B) Lower	Upper
lenders3	1.244	.557	4.982	1	.026	3.470	1.164	10.346
gender3(1)	-.587	.565	1.080	1	.299	.556	.184	1.682
enroll3			.740	2	.691			
enroll3(1)	.978	1.137	.740	1	.390	2.659	.286	24.699
enroll3(2)	-18.588	13223.188	.000	1	.999	.000	.000	.
birthyear	-.006	.036	.030	1	.862	.994	.927	1.066
totaldebt	.000	.000	.674	1	.412	1.000	1.000	1.000
Constant	8.779	70.605	.015	1	.901	6496.565		

a. Variable(s) entered on step 1: lenders3, gender3, enroll3, birthyear, totaldebt.

APPENDIX G

Sample Two Outputs and Data Tables

APPENDIX G

Sample Two Outputs and Data Tables

Table G1

Case Processing Summary for Sample Two

Unweighted Cases[a]		N	Percent
Selected cases	Included in analysis	120	100.0
	Missing cases	0	.0
	Total	120	100.0
Unselected cases		0	.0
	Total	120	100.0

a. If weight is in effect, see classification table for the total number of cases.

Table G2

Dependent Variable Encoding

	Original Value	Internal Value
Dimension0	D	0
	B	1

Table G3

Categorical Variables Coding

			Parameter Coding	
Variable	Code	Frequency	(1)	(2)
enroll 3	G	7	.000	.000
	W	100	1.000	.000
	O	13	.000	1.000
gender 3	M	41	.000	
	F	79	1.000	

Block 0: Beginning Block

Table G4

Classification Table[a,b]

Observed			Predicted		
			Borrow Stat		Percentage
			D	B	Correct
Step 0	Borrow stat	D	103	0	100.0
		B	17	0	.0
	Overall percentage				85.8

a. The constant is included in the model.
b. The cut value is .500

Table G5

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-1.802	.262	47.357	1	.000	.165

Table G6

Variables Not in the Equation[a]

			Score	df	Sig.
Step 0	Variables	gender3(1)	1.464	1	.226
		enroll3	3.961	2	.138
		enroll3(1)	3.961	1	.047
		enroll3(2)	2.406	1	.121
		lenders3	1.595	1	.207
		Birthyear	.126	1	.723
		Totaldebt	.075	1	.784

a. Residual Chi-Squares are not computed because of redundancies.

Block 1: Method = Enter

Table G7

Sample Two Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	11.201	6	.082
	Block	11.201	6	.082
	Model	11.201	6	.082

Table G8

Sample Two Model Summary

Step	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2
1	86.714a	.089	.160

a. Estimation terminated at iteration number 20 because maximum iterations have been reached. The final solution cannot be found.

Table G9

Sample Two Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	4.804	8	.778

Table G10

Sample Two Contingency Table for Hosmer and Lemeshow Test

		Borrow Stat = D		Borrow Stat = B		
		Observed	Expected	Observed	Expected	Total
Step 1	1	12	12.000	0	.000	12
	2	12	11.861	0	.139	12
	3	10	11.060	2	.940	12
	4	11	10.683	1	1.317	12
	5	11	10.540	1	1.460	12
	6	10	10.296	2	1.704	12
	7	12	9.912	0	2.088	12
	8	10	10.122	3	2.878	13
	9	9	9.671	4	3.329	13
	10	6	6.855	4	3.145	10

Table G11

Sample Two Classification Table[a]

Observed			Predicted		
			Borrow Stat		Percentage
			D	B	Correct
Step 1	Borrow stat	D	103	0	100.0
		B	17	0	.0
	Overall percentage				85.8

a. The cut value is .500

Table G12

Sample Two Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95% C.I. for EXP(B)	
								Lower	Upper
Step 1a	gender3(1)	-.693	.615	1.272	1	.259	.500	.150	1.668
	enroll3			.000	2	1.000			
	enroll3(1)	19.814	14867.030	.000	1	.999	4.028E8	.000	
	enroll3(2)	.280	18358.236	.000	1	1.000	1.323	.000	
	lenders3	-1.009	.699	2.084	1	.149	.365	.093	1.435
	birthyear	.017	.043	.150	1	.698	1.017	.934	1.107
	totaldebt	.000	.000	2.113	1	.146	1.000	1.000	1.000
	Constant	-53.678	14867.279	.000	1	.997	.000		

a. Variable(s) entered on step 1: gender3, enroll3, lenders3, birthyear, totaldebt.